The limitations of active embodiment in preschool classrooms

Evaluating the effectiveness of multimodality in a classroom-based narrative intervention

Júlia Florit-Pons[1], Alfonso Igualada[2], and Pilar Prieto[1,3]
[1] Universitat Pompeu Fabra | [2] Universitat Oberta de Catalunya |
[3] Institució Catalana de Recerca i Estudis Avançats

This study assesses the MultiModal Narrative (MMN) intervention aimed at enhancing children's oral narrative skills, focusing on its multimodal component, which includes story enactment by teachers and children, and an audiovisual multimodal narration by a storyteller. A cluster randomized trial was conducted (nine classrooms, 115 children), comparing a control group, a multimodal narrative-based group (complete MMN), and a narrative-based group (MMN without controlling for multimodality). The children's oral narrative macrostructure, microstructure, and perspective-taking skills were measured pre- and post-intervention. Results revealed that only the narrative-based group improved in narrative macrostructure and microstructure skills. No significant effects were found for the multimodal narrative-based group. Further analyses indicated that the children's active embodiment involved in the multimodal component introduced changes in classroom dynamics, thereby constraining children's narrative learning opportunities. The study suggests that multimodality should be integrated in a way that complements, and does not affect, the narrative intervention activities.

Keywords: multimodality, narrative intervention, narrative macrostructure, narrative perspective-taking, narrative microstructure, preschool education

https://doi.org/10.1075/itl.26005.flo | Published online: 19 November 2026
ITL - International Journal of Applied Linguistics ISSN 0019-0829 | E-ISSN 1783-1490
Available under the CC BY 4.0 license. © John Benjamins Publishing Company

Introduction

Narrative development and the need for narrative interventions

The development of the ability to narrate a story typically takes place during the preschool years and continues to progress throughout childhood, constituting one of the more complex abilities in child language development. Early narrative skills are considered as a reliable and ecologically valid measure for evaluating children's linguistic development (for a review, see Dickinson & McCabe, 2001). Around the age of five, children begin to produce coherent narratives which incorporate key story elements such as characters, settings, initiating events, attempts, and outcomes. In addition, narrative skills have also been found to be both predictive and precursors of subsequent linguistic, literacy, social, and academic performance (e.g., Babayiğit et al., 2021; Dickinson & McCabe, 2001). However, substantial individual variability remains in story organization, and children naturally vary in their narrative abilities more broadly. This variability appears particularly consequential for children with, or at risk for, language and learning difficulties, for whom narrative deficits at preschool age tend to persist into later childhood and adolescence if unaddressed (e.g., Botting, 2002). This highlights a concrete need for narrative interventions to support language learning across different learners.

Narrative interventions allow language to be trained holistically, and researchers and practitioners have been devoted to developing interventions that promote oral narrative skills in preschool- and school-aged children. In the last years, different narrative-based interventions have been implemented and have generally been found to be effective (see Nelson & Khan, 2019 and Pico et al., 2021 for reviews focusing on narrative interventions; for a review on more general language-based interventions, refer to Donolato et al., 2023).

Narrative intervention outcomes: Narrative macrostructure, microstructure and perspective-taking

Within the literature on narrative interventions, most studies have focused on boosting children's *narrative macrostructure* (i.e., the structure of the narrative discourse in terms of story elements, such as character, problem, attempt, and resolution; Favot et al., 2021) and *narrative microstructure* (i.e., word- and sentence-level linguistic features used to convey story content, typically evaluated in terms of productivity and complexity, such as word count and lexical diversity; Justice et al., 2006). Despite the variability in the duration of each narrative intervention (ranging from two-session interventions to whole academic year interventions),

the vast majority of studies have reported positive outcomes after implementing the intervention for narrative macrostructure (e.g., Henry & Solari, 2020; Hettiarachchi, 2022), showing that training story elements can also help improve narrative microstructure (e.g., Gillam et al., 2014; Spencer et al., 2018). Indeed, a recent meta-analysis found that narrative interventions have a positive, medium-sized effect on both narrative production ($d = 0.54$) and narrative comprehension ($d = 0.51$) outcomes across studies (Pico et al., 2021). Although narrative macrostructure and microstructure are often examined together, given that story elements are conveyed through words and sentences, these two dimensions are not necessarily fully interdependent: children can, for instance, produce a structurally complete story using relatively simple language, or, conversely, produce linguistically rich language without a coherent story structure. This suggests that narrative macrostructure and microstructure, although related, may draw on partially distinct underlying skills (Bolk et al., 2026).

These studies have also revealed that narrative interventions can promote generalization to other linguistic skills beyond narrative macrostructure and microstructure, including vocabulary knowledge — as a broader language outcome distinct from narrative microstructure (Gillam et al., 2014; Suggate et al., 2021), or even reading and writing (e.g., Spencer & Petersen, 2018; Petersen et al., 2022). For instance, Spencer et al. (2018) implemented the registered narrative intervention *Story Champs* in three preschool classrooms throughout the school year and compared it to three other control classrooms. The authors conducted two narrative assessments in winter and spring and found that, at both time points, the intervention group achieved significantly greater advancement in narrative macrostructure skills (measured with a narrative retelling task) than the control group. For other studies reporting similar findings, see Daelman et al. (2024), Gillam et al. (2014), and West et al. (2021). For similar findings from lab-based studies, see Vilà-Giménez et al. (2019) and Vilà-Giménez & Prieto (2020).

When telling or retelling a story, it is not only relevant to talk about what happened in the story, but also about how the characters felt and thought (Veneziano et al., 2020). Despite the significant effects on narrative measures, few studies have systematically trained and evaluated *narrative perspective-taking* (i.e., the characters' emotions, perspectives, and mental states) but have generally considered emotions as another story element (e.g., Gillam, Gillam & Laing, 2014; Spencer & Petersen, 2012). One exception is the study by Veneziano et al. (2020), which implemented a short, individualized narrative intervention with 5- to 8-year-old French students. In this training session, children were asked to tell a story by looking at a set of images, and then the interventionist started a conversation to talk about the character's mental states. The results showed that after the intervention, the narrative discourses of children receiving the intervention contained

more mental states than those of children in the control group (see also Dodd et al., 2011 for convergent findings in students with autism, suggesting this may be a more general mechanism rather than one specific to typically developing populations). Within the context of preschool classrooms, Pronina et al. (2021) showed that a narrative-based intervention focusing on characters' emotions and mental states was effective in boosting expressive pragmatic skills (in terms of the appropriateness of answering different speech acts). These studies highlight the importance of incorporating direct instruction on emotions and perspectives into narrative interventions.

Educational strategies in narrative interventions

The effectiveness of most narrative interventions can be explained through the use of various educational strategies to support language learning. As reported in the recent review by Pico et al. (2021), most interventions provide specific instructions on unfolding narrative macrostructural elements, typically accompanied by visual support (e.g., story icons, pictures, or storyboards). Another key strategy for training macrostructure elements is to employ verbal support, such as verbal prompting through question-and-answer sequences or positive feedback (e.g., repetitions and expansions). A meta-analysis by Pesco and Gagné (2017) found that verbal scaffolding, alone or in combination with other strategies, was the predominant teaching approach across narrative interventions in early childhood settings, yielding medium average effects on both narrative production ($d = .50$) and narrative comprehension ($d = .58$). In line with these strategies, a recent study by Spencer and Petersen (2020) postulated a set of principles for effective narrative intervention, which included guidelines such as promoting active oral participation and making the intervention as fun as possible, framing the stories in a socially relevant context for children, and providing different opportunities for children to tell and retell the whole story, which in turn allows for generalization to other stories. These interactive, dialogic strategies closely parallel those examined in the shared book reading literature, where similar techniques, such as verbal prompting, open-ended questioning, and elaborative feedback, have been extensively studied as tools to support children's language and narrative development (for a meta-analysis, see Noble et al., 2019). Notably, dialogic approaches to shared book reading have also been shown to support children's broader socioemotional and pragmatic development, particularly when adult-child interactions explicitly incorporate discussion of characters' emotions and mental states (for a review, see Schapira & Grazzani, 2025). Nevertheless, to the best of our knowledge, there is an educational strategy that has been incorporated less frequently into narrative interventions, namely multimodality.

Multimodality in language development and learning: An embodied cognition perspective

Multimodality refers to how we use our body, hands, face, and voice to communicate (Perniss, 2018). A growing body of research has shown that multimodal communicative behaviors, such as gestures and facial expressions, play a foundational role in language development. Early multimodal skills are predictive of later oral language, vocabulary, and narrative competence, suggesting that children's multimodal abilities scaffold their acquisition and organization of language (for reviews, see Kirk et al., 2022; Vilà-Giménez & Prieto, 2021). Understanding these links provides a rationale for incorporating multimodal strategies into narrative interventions, as they may offer an additional pathway for supporting language learning.

The theoretical rationale for why multimodality might support narrative learning comes from the embodied cognition framework, which posits that cognitive and linguistic processes are grounded in sensorimotor experience (Adams, 2010; Foglia & Wilson, 2013). From this perspective, several distinct but complementary mechanisms have been proposed to explain how gestures and bodily enactment can support language processing and learning. First, gestures can reduce the cognitive load associated with speech production and comprehension by freeing up cognitive resources that can be allocated to linguistic processing, such as lexical retrieval and sentence formulation (Hostetter & Alibali, 2008). Second, observing or producing gestures and physical actions can strengthen the encoding of semantic and event-based information by creating richer, motorically grounded memory traces, consistent with the well-documented "enactment effect" in memory research, whereby self-performed or observed physical actions are remembered better than purely verbal information (for a review, see Roberts et al., 2022). Third, whole-body enactment (i.e., embodiment), in particular, may support perspective-taking and mental-state understanding by allowing children to physically engage in embodied simulations of characters' actions and emotional states (Glenberg, 2010). Taken together, these mechanisms provide a theoretical basis for expecting multimodal strategies to support narrative learning across its macrostructural, microstructural, and perspective-taking dimensions. Finally, the embodied cognition framework aligns with the multimodal enrichment paradigm (Mathias & von Kriegstein, 2023), which claims that our body can promote language learning (see also Dargue et al., 2019; Rohlfing, 2019; Vilà-Giménez & Prieto, 2021 for reviews). Building on this theoretical rationale, in this paper we adopt this embodied conceptualization of multimodality. The following section considers how multimodal, embodied strategies have been operationalized as educational tools within narrative interventions specifically.

Multimodality as an educational strategy in narrative intervention

In the context of narrative storytelling, multimodal cues can serve several distinct communicative functions. Iconic gestures and bodily movements can represent concrete or abstract concepts, such as actions or events (e.g., Demir et al., 2015; Macoun & Sweller, 2016); deictic gestures can direct attention to referents in the narrative (e.g., pointing to a character or object); non-referential (beat) gestures can highlight prosodic prominence or structure discourse (McNeill, 1992; Prieto et al., 2018; Vilà-Giménez et al., 2019); and facial expressions and vocal prosody can convey emotional valence or affective state. Within this broader set of functions, enactment refers more specifically to the physical embodiment of a character's actions, emotions, or perspective, for instance by a storyteller narrating and acting out the story (e.g., Vilà-Giménez et al., 2019), or by children themselves being encouraged to gesture or physically act out story events (e.g., Bernstein et al., 2024; Vilà-Giménez & Prieto, 2020). Importantly, children can be exposed to these multimodal cues through *perception* (e.g., observing a storyteller gesturing or enacting the story) or through *production* (e.g., enacting the story themselves), or both.

Lab-based studies have found that observing gestures during a storyteller's narration (i.e., gesture perception) enhanced children's narrative recall, comprehension and telling skills. This has been shown both for iconic and representational gestures depicting the actions and events of the story (e.g., Guilbert et al., 2021; Macoun & Sweller, 2016) and for non-referential beat gestures (Vilà-Giménez et al., 2019). This facilitative effect is thought to be mediated by gestures' impact on children's attention and memory encoding: iconic gestures are believed to provide a semantic cue that strengthens memory for story content, while non-referential beat gestures are thought to operate by directing attention toward prominent narrative information (Aussems & Kita, 2019; So et al., 2012). Interestingly, Vilà-Giménez and Prieto (2020) also showed that both observing and producing non-referential beat gestures (i.e., gesture perception and production) was effective for promoting five-year-old children's narrative macrostructure skills. However, it should be noted that these abovementioned lab-based studies tend to focus mainly on hand gestures, with less emphasis on broader forms of multimodality involving other articulators.

Based on this, a handful of recent studies have gone one step further and have tested the effects of multimodality in real-life classrooms, incorporating not only hand gestures but also broader forms of enactment such as active embodiment and dramatization (e.g., Bernstein et al., 2024; Nicolopoulou et al., 2015). Actually, Pesco and Gagné's (2017) meta-analysis of narrative instruction in early childhood settings found that combining verbal strategies with nonverbal ones, such

as engaging children in enacting stories, increased effects on narrative skills relative to verbal strategies alone. However, this meta-analysis was not specifically focused on multimodality, and only a small subset of the reviewed studies incorporated such strategies. To the best of our knowledge, only a few classroom-based narrative interventions have systematically incorporated a multimodal component, with mixed findings depending on the outcome measure targeted. Regarding narrative-specific outcomes, one example is the *Story Champs* intervention (Spencer & Petersen, 2012) for preschool- and school-aged children, in which the interventionist uses conventional or metaphoric gestures to represent abstract concepts, such as the different macrostructural elements (e.g., the conventional "thumbs down" gesture to represent the problem of the story). Although the specific contribution of these gestures within *Story Champs* has not been tested, children receiving the intervention have shown greater gains than comparison groups across a range of narrative measures, including story (re)telling, story-based language comprehension, personal story generation, and narrative writing (e.g., Petersen et al., 2022; Spencer et al., 2018). More broadly, other studies have incorporated story enactment strategies as drama-based activities using pantomime gestures, which convey a narrative or story in the absence of speech (see e.g., Bernstein et al., 2024; see also Corbett et al., 2016 for autistic children and Nicolopoulou et al., 2015 for children from households with a low income, suggesting the potential applicability of these strategies across diverse populations). For instance, Bernstein and colleagues (2024) compared children's narrative recall abilities after receiving a classroom-based storybook reading intervention delivered either with or without a drama-based activity in which children physically enacted the main character (i.e., gesture production). Their results showed that, although there were no differences between groups in the quality of children's story retell (i.e., narrative macrostructure, assessed based on the introduction and organization of story elements), those receiving the intervention used more gestures when retelling the stories. In addition, Ionescu and Ilie (2018) showed that a group of preschool children exposed to an embodied narrative intervention, in which they themselves dramatized story events using story-related objects, showed greater performance in retelling the story and in remembering key vocabulary and idioms than the group of children that had been exposed to only listening to the story.

Other studies have tested the intervention effects on broader socioemotional and pragmatic outcomes, involving child-directed storytelling and story enactment. For instance, Pronina et al. (2021) incorporated multimodality into a narrative intervention through the use of facial expressions by both the interventionist and the children themselves (i.e., gesture perception and production) to depict different emotional and mental states, comparing a control group, a

group receiving the intervention with this systematic multimodal embodiment, and a group receiving the same intervention without it. The results showed that both groups receiving the intervention (with no difference between them) improved their expressive pragmatic skills (measured through a speech act production task), but that the intervention was not effective in improving emotion understanding and mental state comprehension. Also, a recent review has found that drama-based interventions can support a range of socioemotional outcomes in children, including the comprehension and expression of emotions of oneself and others, social awareness, self-regulation, working memory and decision-making, although effects vary depending on the type, intensity, and implementation fidelity of the intervention (Frydman & Mayor, 2024).

Overall, with all these studies, we have seen the value of oral narrative interventions for children's language development. However, the specific contribution of multimodality to narrative intervention practice is less clear. Although studies conducted in an individualized or lab-based context have reported positive effects of multimodality, the only two studies testing the effects of multimodality with groups of typically developing students in classrooms (e.g., Bernstein et al., 2024, Pronina et al., 2021) have shown null results. Several methodological differences between these two sets of studies may help explain this discrepancy. Lab-based studies were typically delivered one-on-one by researchers under tightly controlled conditions and isolated a single, discrete gesture type (e.g., iconic or non-referential beat gestures) during a brief exposure, whereas the classroom-based studies were delivered to groups of children by teachers in more naturalistic settings, incorporating broader and more complex forms of multimodality (e.g., embodiment, dramatization) that are harder to implement consistently and to isolate from other intervention components. The two sets of studies also differ in their outcome measures and timing of assessment: lab-based studies primarily assessed immediate recall following a single exposure, whereas classroom-based studies assessed broader, more distal narrative or socioemotional outcomes, measured not immediately after the intervention but on a separate day, both of which may reduce sensitivity to the specific contribution of multimodal input. Taken together, these differences suggest that the null classroom-based findings may reflect challenges in implementation and measurement, rather than indicating that multimodality is ineffective in applied settings.

To address the conflicting evidence between the positive effects of multimodality found at the more individual one-to-one lab-based settings and the lack of effects at the classroom level, we believe that further research is needed to rigorously evaluate the effectiveness of this multimodal component in real-world classroom contexts, where findings could inform broader educational practice. Moreover, provided that teachers inherently use gestures in their professional

practice, we need more research on how multimodality should be systematically incorporated into narrative intervention practice, and its potential effectiveness should be further investigated. This approach would enable teachers to not only implement this strategy effectively in the classroom but also acknowledge its potential value (e.g., Seccia & Goldin-Meadow, 2024).

The present study

The present study aims to assess the effectiveness of the novel multimodal narrative-based intervention (the MultiModal Narrative (MMN) intervention program; Florit-Pons, Igualada & Prieto, 2025), and particularly, the potential effectiveness of its core multimodal component, by testing whether the MMN intervention is effective in improving preschool children's oral narrative skills (i.e., macrostructure, microstructure, and perspective-taking). The main novelty of the MMN intervention lies in its incorporation of this multimodal component, a form of embodiment in which narratives are enacted through coordinated speech, hand gestures, facial expressions and bodily movements. This multimodal component is delivered through three social actors involved in the intervention: a video-recorded storyteller, who uses iconic and conventional gestures and facial expressions to represent story actions, emotions, and characters' perspectives, as well as non-referential gestures to structure the discourse and highlight important information (i.e., gesture perception); the classroom teacher, who similarly uses gestures and facial expressions to represent these same story elements (i.e., gesture perception); and the children themselves, who are encouraged to physically and actively embody these story elements (i.e., gesture production; see the Intervention phase subsection for more details). We hypothesize that children who receive the intervention with a systematic multimodal component will show greater improvements than those who receive the same narrative-based intervention without the controlled multimodal component and those that do not receive the intervention. This hypothesis builds on the embodied cognition mechanisms outlined above: we expect that engaging with the narrative through the body, rather than through oral language activities alone, will support children's narrative skills by strengthening memory encoding of story events, by reducing the cognitive load associated with language formulation, and by allowing children to physically embody and thus more readily access characters' emotions and mental states. More broadly, we expect that multimodal strategies will allow children to experience the events of the story, helping them to internalize and 'live' the narrative from a first-person perspective.

Method

Design

To test the effectiveness of the MMN intervention, particularly the contribution of its multimodal component, this study followed a cluster randomized controlled trial with three different groups: a control group that did not receive the intervention and two experimental groups. One experimental group received the MMN intervention (henceforth multimodal narrative-based group), and another one received the MMN intervention without the multimodal component (henceforth narrative-based group). The study involved three participating schools, each contributing three classrooms. Group assignment was done at the classroom level (i.e., entire classrooms within each school, rather than individual students, were assigned to groups), with each group consisting of three classrooms that were randomly allocated to one of the three groups.

Participants

A total of 115 children (55 girls) from an initially recruited sample of 136 children participated in this study (see Appendix I for the CONSORT 2010 flow diagram of the process from participant enrollment until analysis). An *a priori* power analysis indicated that 111 participants were required to detect small effects ($\alpha = .05$, power = .80).

Participants were Spanish-Catalan bilinguals[1] aged between 4 years 9 months and 6 years 2 months attending their last year of preschool. This age range was selected for two main reasons. First, from a narrative development perspective, it is around age 5–6 that children begin producing "true narratives" (Applebee, 1978); that is, complete discourses that include all story elements, are structured in a logical and causally/temporally connected way, and incorporate perspective-taking elements such as characters' motivations and emotions. Second, from a gestural development perspective, this is also the age range in which children begin to produce a broader range of gesture types during narrative tasks, moving beyond iconic gestures that visually depict information to include more complex, adult-like forms such as non-referential gestures (Mathew et al., 2018; Rohrer et al., 2022). This age range therefore allowed us to include children who were

1. Although all children were bilingual, Catalan is the vehicular language and the language of instruction in their schools. Therefore, all assessment and intervention activities were carried out in Catalan, and bilingual status was not a variable of interest in the present study or a factor in the study design.

developmentally positioned both to benefit from a narrative-based intervention and to produce different gestures targeted by the multimodal condition.

Participants were recruited from nine preschool classes in three public schools in L'Hospitalet de Llobregat, a city in the metropolitan area of Barcelona, Catalonia. All three schools were located in the same neighborhood, with middle-low to middle socioeconomic status (according to the socioeconomic territorial index calculated by the Statistical Institute of Catalonia). The inclusion criteria were as follows: children attending the last year of preschool (I5 in the Catalan education system), with the only restriction that they needed to have normal hearing. The final sample consisted of 106 children with typical development, while nine children had a professional report indicating developmental difficulties with language and communication. Participants' structural language abilities and non-verbal IQ were evaluated using standardized tests before the intervention (K-BIT by Kaufman & Kaufman, 1990, for non-verbal IQ; and CELF-5 by Wiig et al., 2013, for structural language). A set of pairwise *t*-tests was performed to assess potential differences between the three groups. No significant differences were found between the groups in terms of age ($ps > .361$), structural language ($ps > .135$), or non-verbal IQ ($ps > .467$). See Table 1 below for the descriptives of the participants' demographic, linguistic and cognitive characteristics.

Table 1. Participants' demographic, linguistic and cognitive characteristics

	Control	Narrative-based	Multimodal narrative-based
N of participants	38 (16 F, 22 M)	41 (20 F, 21 M)	36 (19 F, 17 M)
N of participants with language difficulties	3 (3 M)	2 (2 M)	4 (2 F, 2 M)
Age: Mean (SD)	5.38 (0.33)	5.40 (0.35)	5.45 (0.38)
Age: Range	4.75–6.08	4.75–6.17	4.75–6.17
CELF-5: Mean (SD)	89.14 (16.99)	92.76 (16.19)	87.29 (13.91)
CELF-5: Range	40–115	55–120	55–113
K-BIT: Mean (SD)	101.58 (12.73)	102.73 (12.46)	100.57 (13.62)
K-BIT: Range	72–123	77–141	66–123

This study received ethical approval from the Institutional Committee for Ethical Review of Projects at the host university of the first and third authors (2021, ref.: 228) and from the Regional Ministry of Education (Catalan Government). Legal guardians were provided with detailed information about the project and provided written consent for their child to participate in the study.

Pre- and post-intervention assessment

Materials

To determine whether the intervention was effective in boosting children's oral narrative skills, children completed a narrative telling task at both pre and post-intervention assessments, consisting of two untrained stories.

The two stories were animated wordless cartoons from the series *Die Sendung mit der Maus*, each lasting approximately 50 seconds. At pre-intervention, children first narrated a story featuring the main character, a mouse, and then a second story featuring two characters: the mouse and a small elephant. At post-intervention, children narrated two new stories featuring the same characters. This ensured the stories were novel at each assessment and could therefore serve as a test of generalization rather than memory or repeated exposure. At both time points, the stories were administered in a fixed order: first the one-character story, and then the two-character story.

Procedure

The task was administered to each child individually in a silent room at the school during school hours. First, the child watched the cartoon, and the experimenter then asked him/her to stand up and narrate the story. The child was asked to stand up with the aim that their gestures and enactment of the story could be further investigated in a follow-up study, and to encourage a more natural use of multimodal expression and gesture. The task was presented in the form of a game, such that after the child narrated the story, the experimenter had to guess — from a set of four still images — the story the child was talking about. Throughout the narration, children received continuous positive reinforcement (e.g., nodding, smiling, verbal feedback such as *mhm* or *aha*) regardless of the content or completeness of their story. If a child stopped narrating before reaching the end of the story, the experimenter provided up to a maximum of three neutral prompts to encourage continuation (e.g., "And what happened next?", "Tell me more", "Anything else?"), and if the child still did not continue, they were reassured that it was fine if they did not remember everything and was asked to share only the parts they did recall (e.g., "It's okay if you don't remember everything, just tell me the parts you do remember"). This process was repeated for the second story. The task was administered before and after the intervention (i.e., between one or two weeks after the last intervention session). All pre- and post-intervention sessions were video recorded for them to be revised and coded after the sessions.

Coding

Each narration was coded for narrative macrostructure, microstructure and perspective-taking. While macrostructure and perspective-taking were coded during the data collection by the experimenter (i.e., a research assistant), narrative microstructure was coded *a posteriori* by the first author together with two other predoctoral researchers and a research assistant. All coders were graduate-level researchers who had received specific training from the first author using sample narratives until full familiarity with the coding schemes was reached, allowing them to easily code for macrostructure and perspective-taking at the same time.

First, narrative macrostructure was coded using a 0–6 coding scheme adapted from Demir et al. (2014) to assess children's introduction of different macrostructural elements, with the score increasing depending on the number of story elements introduced. For instance, a score of 6 indicated that the child introduced all elements and added extra details about what happened during the plot, whereas a score of 0 indicated that the child did not introduce any elements or a descriptive sequence (see Appendix II).

Narrative microstructure involved two narrative productivity measures, namely the total number of words (TNW, also known as *tokens*) and the number of different words (NDW, also known as *types*). After all narrations were transcribed following the CHAT guidelines (MacWhinney, 2000), TNW and NDW were automatically calculated using the CLAN software (MacWhinney, 2000). Transcription began once the child started narrating and any prior interaction (e.g., task instructions) was not transcribed. Non-narrative interactions occurring during the narration (e.g., meta-comments, off-topic reflections) were transcribed but excluded from the analysis. The coding for narrative microstructure could only be performed on 107 participants (rather than all 115), provided that eight participants (three from the control group, two from the narrative-based group and three from the multimodal narrative group) had to be excluded because the narration could not be video recorded due to technical difficulties or because there was too much background noise, and therefore the narration could not be annotated.

Narrative perspective-taking was coded using a 0–6 coding scheme adapted from Dodd et al. (2011), which assessed children's ability to introduce emotional and mental terms within the narration (e.g., *angry, happy, sad, surprised*, and *thinking, willing, realizing*), with the score increasing depending on the number of elements introduced as well as the relation among them. A score of 6 corresponded to a narration that introduced at least two emotional terms, the cause of these emotions, and at least two mental terms, while a score of 0 corresponded to a narration that did not include any emotional or mental term (see Appendix III).

An example of a target story, its expected macrostructural and perspective-taking elements, and a sample child narration illustrating how the scoring scheme was applied is presented in Appendix IV. Partial or approximate responses were coded as correct when they preserved the functional role of the element within the story's plot, even if specific details differed from the original version (e.g., a child describing the mouse as a bear was still credited with introducing the corresponding character, as long as its narrative function remained consistent with the story structure).

Inter-rater reliability

Provided that narrative macrostructure and narrative perspective-taking codings were perceptual assessments, inter-rater reliability was calculated using data from 20 participants, corresponding to 80 stories (2 stories × 2 time points) and 17.39% of the data. The participants were randomly selected from the whole set of participants. Reliability was calculated using the original scores and the scores of a trained research assistant. Cohen's kappa (weighted kappa) was used to calculate inter-rater reliability. The results for narrative macrostructure showed 83.8% agreement among coders and a kappa of .899. Regarding narrative perspective-taking, the results indicated 97.5% agreement and a kappa of .861.

Intervention phase

The MultiModal Narrative (MMN) intervention

The MMN intervention is a multi-tiered narrative-based intervention developed in a co-creation fashion with more than 90 preschool teachers and speech-language therapists working in Catalan schools and speech therapy centers (Florit-Pons, Prieto & Igualada, 2025). In this study, we used the universal tier (tier 1) of the MMN intervention, which was implemented at the classroom level.

Intervention protocol

The MMN intervention aims to train oral narrative skills through interaction activities and involves a dedicated multimodal component. The intervention protocol revolves around three different wordless cartoons from the series *Chigüiro*, each of which is trained for three consecutive days (three stories × three sessions = nine intervention sessions). The protocol is as follows. All intervention sessions follow the same initial set of activities: introduction to the session, watching the wordless cartoons, a short teacher-children interaction about the story plot, and watching a video of the storyteller narrating the story (without the animated cartoons, see Appendix V for the scripts). After these activities, each session has a central activity. In the **first session** of each story, the central activity consists of

a sequential enacted retelling using a question-and-answer sequence. First, the teacher asks a question about each story element, the children answer and she then gives feedback. While providing this model, the teacher enacts the action and after that, encourages all children to stand up and embody it. This process is repeated for all elements of the story. In the **second session** of each story, the central activity centers around the retelling and dramatization of the story. While children remain seated, the teacher asks a question targeting all story elements. For each question, two children are invited to the front of the class: one takes on the role of the storyteller and responds to the question, while the other takes on the role of the actor and embodies the corresponding action or emotion. The teacher then provides positive feedback after the performance. This process is repeated for all story elements with different pairs of children, with all pairs participating at least once. Once all questions have been answered, children are invited to volunteer to retell and dramatize the entire story in pairs in front of the class. The central activity during the **third session** of each story focuses on the generation of personal stories. First, the teacher tells and enacts a personal narrative related to the trained fictional story and encourages children to stand up in front of their classmates, tell their personal stories and embody them. For more detailed information on the MMN intervention protocol, please refer to Florit-Pons, Igualada & Prieto (2025). All the intervention materials and scripts are openly available at the Open Science Framework.

The intervention was implemented over three consecutive weeks, involving three weekly sessions, each of which lasted approximately 25 to 30 minutes. All intervention sessions were implemented in Catalan, the vehicular language in schools in Catalonia, and by each classroom teacher, for it to be as natural as possible for children receiving the sessions.

Differences between the two experimental groups

As mentioned above, two different groups received the same MMN intervention, with the exception that one group received it with the presence of the multimodal component (multimodal narrative-based), the other received it without the multimodal component (narrative-based).

The multimodal component is integrated into the intervention involving the three social actors of the intervention, namely, the audiovisual recordings of the storyteller, the teacher, and the children. First, we have the video of the storyteller explaining the story: children in the multimodal narrative-based group watched the video of the storyteller (i.e., gesture perception) in which she used not only child-directed prosody but also child-directed enactment strategies, implemented through the use of different gestures and facial expressions. Specifically, the storyteller's gestures were designed to correspond to the different target macrostructural

and perspective-taking elements of the story: iconic gestures (involving either the hands or the whole body) were used to represent the main actions corresponding to the key story elements, facial expressions were used to convey characters' emotions, and conventional gestures (e.g., a thinking gesture) were used to represent characters' perspectives and mental states. In addition, non-referential gestures were used to structure the discourse, typically accompanying key connectors (e.g., "and then", "in the end"). The second social actor is the teacher, who apart from using natural multimodal strategies during the intervention, received scripted prompts to embody the main macrostructural and perspective-taking elements of the stories (i.e., gesture perception). Finally, the third actor refers to the children receiving the intervention: teachers also had instructions to encourage the children to stand up and also actively embody the macrostructural and perspective-taking elements of the stories themselves (i.e., gesture production). This embodiment, both by the teacher and by the children themselves, mainly involved the use of iconic gestures (hand and whole-body) and facial expressions, serving the same representational purpose as described for the storyteller (i.e., depicting story actions and characters' emotions and perspectives). As for the narrative-based group, the video of the storyteller included a narration in which the storyteller was naturally telling the story without using gestures but maintaining natural child-directed facial expressions and a child-directed prosodic style. In addition, teachers in the narrative-based group did not have any specific instruction to enact the stories but were rather asked to act naturally. Children were not encouraged to embody the story and thus remained seated throughout the session. Figure 1 shows a visual representation of the differences between the two experimental groups.

It is important to note that, given the naturalistic, classroom-based nature of this intervention, our aim was not to create a fully "non-multimodal" condition devoid of any gesture, as this would constitute a highly artificial and ecologically invalid communicative context. Indeed, gesture is considered a ubiquitous and largely spontaneous feature of human communication, including in classroom settings, occurring even when speakers are not explicitly instructed to gesture (Macoun & Sweller, 2016). However, although teachers spontaneously gesture in the classroom, it remains unclear whether they do so in the most effective way for supporting children's learning, and explicit instruction on gesture use may be necessary to harness its full pedagogical potential (Hostetter et al., 2006; Macoun & Sweller, 2016). The multimodal narrative-based condition in the present study builds on this idea by providing teachers with specific, targeted instructions on gesture use, allowing us to examine how such gestural input and embodiment function in real classroom contexts, above and beyond the spontaneous gesture already naturally present in teacher-child interaction. Suppressing this natural,

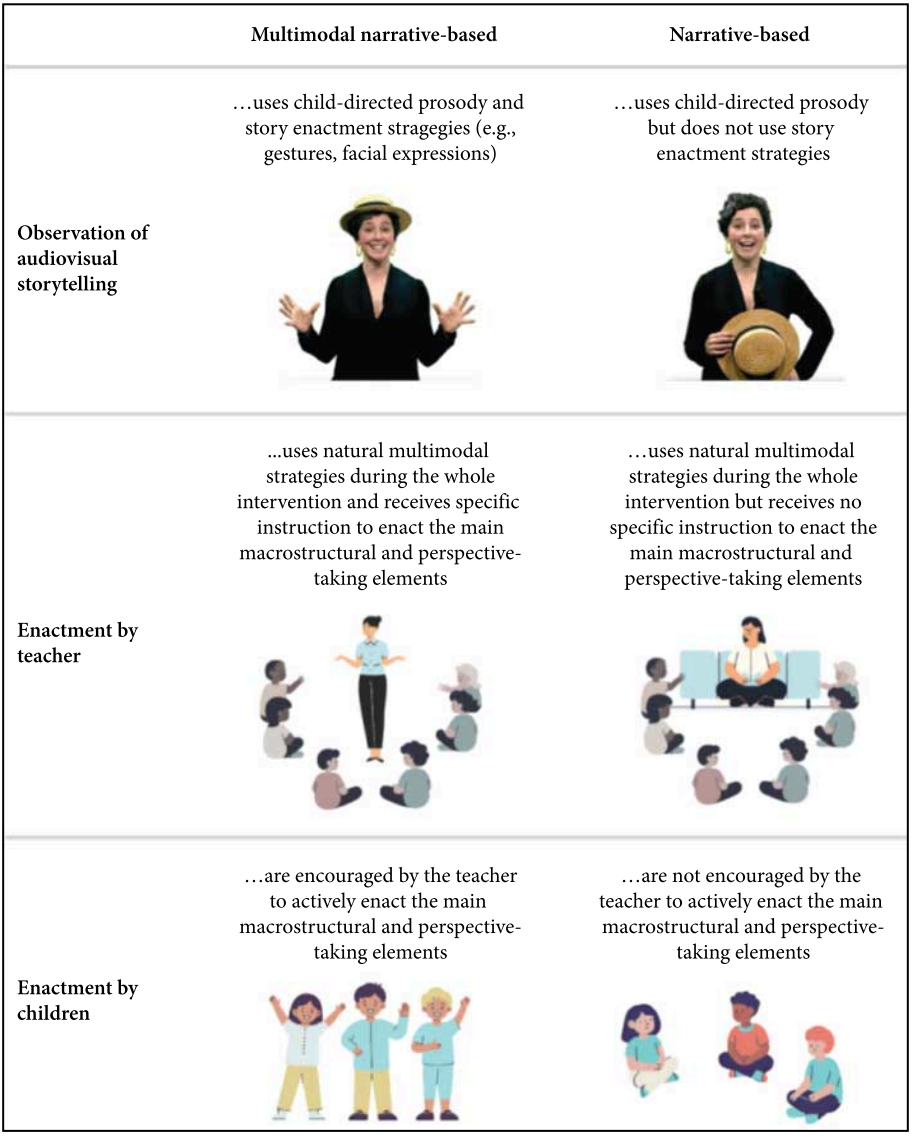


Figure 1. Schematic representation of the implementation of the multimodal component. Left panel shows the multimodal narrative-based group and the right panel shows the narrative-based group

spontaneous gesture entirely in the narrative-based condition would therefore not only be exceedingly difficult to enforce, but would also compromise the ecological validity of the intervention and obscure the specific contribution of gesture instruction, which is the focus of the present study. Instead, the two conditions

were designed to differ specifically in the presence versus absence of this systematic, targeted multimodal input and enactment, as detailed below.

To ensure that the videos of the storyteller were comparable across the two groups, a group of eight preschool teachers (unfamiliar with the intervention and blind to the intervention group) evaluated the videos on a 1–7 Likert scale. It was found that both stimuli were similar in terms of how natural the narration was (narrative-based: $M = 5.78$, $SD = 1.12$; multimodal narrative-based: $M = 6.44$, $SD = 0.85$).

Intervention fidelity

The intervention fidelity was assessed using two complementary measures, e.g. teachers' self-assessment and external *a posteriori* evaluation by a research assistant. First, after each intervention session, all teachers were asked to complete a short questionnaire that reported the duration of the session and whether they had followed the intervention protocol. The average duration of the sessions was 27.93 minutes ($SD = 4.87$), with a similar duration for both groups (narrative-based: $M = 28.12$, $SD = 6.10$; multimodal narrative-based: $M = 27.78$, $SD = 3.86$). Teachers reported that, in response to a yes/no item ("In today's session, did you follow the intervention protocol?"), they answered "yes" 95% of the time (narrative-based: 91.30%; multimodal narrative-based: 100%). In cases where teachers responded "no", they were given space to briefly describe what they had changed or deviated from in the protocol.

Additionally, intervention fidelity was evaluated *a posteriori* through video recordings of the intervention sessions. A research assistant blinded to the intervention group watched a total of 18 intervention sessions (i.e., three sessions × six teachers), corresponding to 33% of the data, and assessed the teacher's treatment fidelity using a 1–7 Likert scale (1 = the teacher did not follow the intervention protocol and 7 = the teacher followed the protocol step by step). The results showed that the interventionists systematically followed the intervention procedure ($M = 6.33$, $SD = 0.84$; narrative-based: $M = 5.67$, $SD = 0.71$; multimodal narrative-based: $M = 7$, $SD = 0$). The coder was also asked to assess the implementation of multimodality by the teachers (i.e., how expressive they were with their body using different bodily gestures and facial expressions) also using a 1–7 Likert scale (1 = the teacher never used multimodal strategies [i.e., hand gestures, facial expressions, story enactment] and 7 = the teacher constantly uses multimodal strategies). Teachers in both groups used gestures and facial expressions throughout the sessions (multimodal narrative-based: $M = 6.56$, $SD = 0.73$; narrative-based: $M = 3.33$, $SD = 1$). This was because all teachers (both in the narrative-based group and the multimodal narrative-based group) were asked to act naturally, with the only difference being that the teachers in the multimodal

group were given specific instructions on how to use gestures throughout the session.

A complementary assessment of the implementation of the intervention was also conducted *a posteriori* with the aim to further investigate potential classroom behaviors across the groups.

Control group: Treatment as usual

Finally, the three class groups constituting the control group did not receive any intervention but continued with their usual sessions involving standard curricular activities. Detailed *a posteriori* reports from the three teachers explained that during the target weeks in which the other groups were taking the interventions they were involved in educational activities to promote new vocabulary and phonological awareness, to establish routines, and to initiate mathematical thinking. No oral narrative activities were performed in any control group.

Statistical analyses

To assess the effectiveness of the intervention on children's oral narrative skills, four Linear Mixed-Effects (LME) models were performed using the *lme4* package (Bates et al., 2015) in R. The score for each narrative measure (i.e., macrostructure, microstructure (TNW and NDW) and perspective-taking), which was the average score of the two narrations, was established as the dependent variable. All models included Test (two levels: pre- and post-intervention) and Group (three levels: control, narrative-based, multimodal narrative-based) as fixed factors, together with their two-way interaction. The best random-effects structure was determined using the *performance* package (Lüdecke et al., 2021). Candidate models included varying intercepts for participants, language score, classrooms, and schools, either alone or in combination. The final models included varying intercepts by language and by participant, or varying intercepts by participant only. Omnibus tests for fixed effects were obtained via Type II Wald chi-square tests (car package; Fox & Weisberg, 2019), and fixed-effect coefficients, standard errors, and 95% Wald confidence intervals were extracted for each model to allow full interpretation of effect direction and magnitude. Degrees of freedom for fixed-effect coefficients were estimated using Satterthwaite's method (lmerTest package; Kuznetsova et al., 2017). Finally, for all significant main effects and interactions, post-hoc pairwise comparisons were performed using the Bonferroni correction with the *emmeans* package (Lenth, 2021), using Kenward-Roger degrees of freedom, and a measure of effect size (via Cohen's *d*). Baseline equivalence between groups on each narrative outcome was additionally assessed using pairwise comparisons of estimated marginal means at the pre-intervention time-

point, from the same models described above. Note that fixed-effect coefficients reflect differences relative to the reference levels of Group (control) and Test (post-intervention), whereas emmeans pairwise comparisons reflect simple effects — i.e., direct group-vs-group or time-vs-time differences at a given level of the other factor — and are therefore the basis for the directional interpretations reported in the Results.

Particularly, the analyses for narrative macrostructure and narrative perspective-taking were performed using data from all 115 participants, whereas the analyses for narrative microstructure were performed with data from 107 participants (see *Coding* subsection for more details).

As a follow-up to the general assessment, we descriptively examined the narrative skills of children identified as having language and communication needs to explore whether they showed improvements after receiving only tier 1 instruction. Given the small sample size ($N = 9$), no inferential statistics were conducted. Instead, we report their individual scores and compare them to the typically developing peers.

Results

This section describes the results obtained from the different analyses. We first provide descriptive statistics for all narrative measures, and then report the statistical results of the LME models conducted with each narrative measure.

Descriptives

Table 2 shows the descriptive statistics for each narrative measure at pre- and post-intervention assessments broken down by group (control, narrative-based, multimodal narrative-based).

Table 2. Descriptives for pre- and post-intervention assessments

		Control		Narrative-based		Multimodal narrative-based	
	Value	**PRE**	**POST**	**PRE**	**POST**	**PRE**	**POST**
MACRO	*M*	3.86	3.64	3.65	4.02	3.96	4.08
	(*SD*)	(1.25)	(1.13)	(1.16)	(0.90)	(0.77)	(0.99)
	Range	0–5.50	0–5.50	0–5.50	2–5.50	2–5.50	1–5.50
MICRO.: TNW	*M*	40.96	36.86	37.88	46.65	42.40	42.74
	(*SD*)	(16.03)	(15.80)	(13.19)	(14.33)	(12.72)	(14.28)
	Range	5–71	18–78	12–64	16–74	18–68	10–65
MICRO.: NDW	*M*	26.03	23.97	24.36	28.28	25.74	25.89
	(*SD*)	(7.73)	(7.13)	(6.70)	(6.74)	(5.88)	(6.29)
	Range	4–40	15–42	10–39	13–42	15–39	9–39
PERS.	*M*	0.17	0.05	1.16	0.90	0.29	0.17
	(*SD*)	(0.37)	(0.16)	(0.25)	(0.12)	(0.35)	(0.40)
	Range	0–1.50	0–0.50	0–1	0–1.50	0–1	0–2

Note 1. MACRO. = macrostructure; MICRO. = microstructure; TNW = total number of words; NDW = number of different words; PERS. = perspective-taking; PRE = pre-intervention; POST = post-intervention

Note 2. The possible range for narrative macrostructure and narrative perspective-taking was 0–6. The possible range for the narrative microstructure measures was 0–NA, with NA indicating that no theoretical maximum value is defined for this measure.

Note 3. The score corresponds to the average score of the two narrations.

Narrative macrostructure

The LME model assessing narrative macrostructure showed no main effects for Test ($\chi^2(1) = 1.06$, $p = .304$) nor Group ($\chi^2(2) = 4.46$, $p = .107$). The two-way interaction between Test and Group was significant ($\chi^2(2) = 7.48$, $p = .024$). Fixed-effect estimates for the model are presented in Table 3. Follow-up pairwise comparisons indicated that only the narrative-based group significantly improved from pre- to post-intervention ($\beta = 0.38$, $SE = 0.15$, $t(111) = 2.53$, $p = .013$, $d = .56$), while the control and multimodal narrative-based groups did not change significantly ($p = .177$ and $p = .597$, respectively). In addition, no significant differences were observed between groups at post-intervention (control vs. multimodal narrative-based: $p = .079$; control vs. narrative-based: $p = .442$; multimodal vs. narrative-based: $p = 1.000$) nor at pre-intervention (control vs. multimodal narrative-based: $p = 1.000$; control vs. narrative-based: $p = .607$; multimodal vs. narrative-based: $p = .077$; see Figure 2).

Table 3. Fixed-effect estimates for the linear mixed model predicting narrative macrostructure

Fixed effect	*β*	*SE*	95% CI	*t*	df	*p*
Intercept	3.592	.189	[3.22, 3.96]	19.06	92.03	<.001
Condition: Multimodal narrative-based	0.507	.224	[0.07, 0.95]	2.26	150.21	.025
Condition: Narrative-based	0.313	.214	[−0.11, 0.73]	1.47	149.13	.145
Test: Pre-intervention	0.211	.155	[−0.09, 0.51]	1.36	111.00	.177
Condition (Multimodal) × Test (Pre)	−0.296	.224	[−0.74, 0.14]	−1.32	111.00	.189
Condition (Narrative-based) × Test (Pre)	−0.589	.215	[−1.01, −0.17]	−2.73	111.00	.007

Note 1. Reference (baseline) levels: Condition = control group; Test = post-intervention.
Note 2. Model fit: R^2 marginal = .035, R^2 conditional = .633, ICC = .62, AIC = 619.3, BIC = 650.2.
Note 3. Random effects: Participant intercept variance = 0.285 (*SD* = 0.534); CELF intercept variance = 0.461 (*SD* = 0.679); residual variance = 0.457 (*SD* = 0.676).
Note 4. CIs computed using the Wald method.

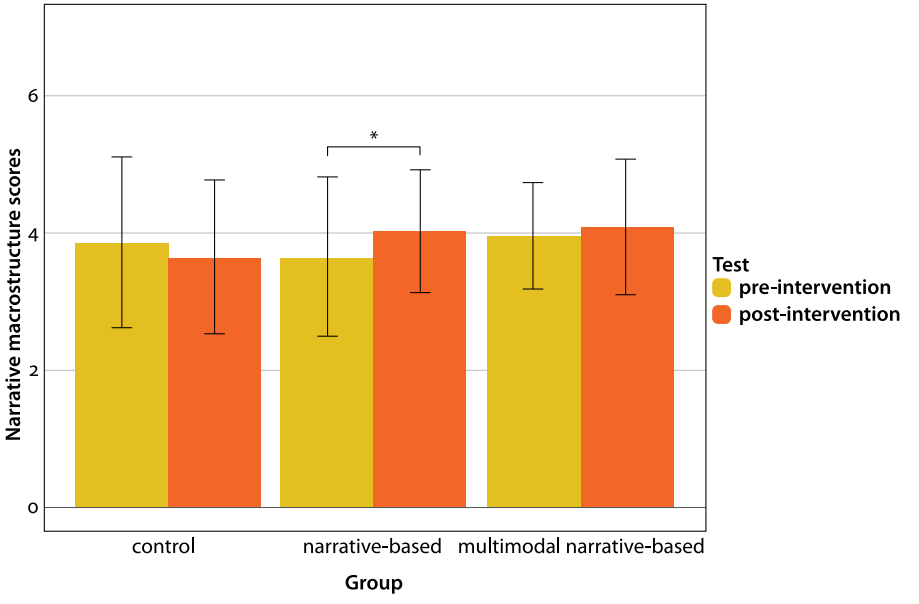


Figure 2. Mean narrative macrostructure scores broken down by Test and Group
Note. * $p \leq .05$. Error bars represent standard deviations.

Narrative microstructure

Total Number of Words (TNW)

The model evaluating TNW reported no main effects for Test ($\chi^2(1) = 1.98$, $p = .160$) nor Group ($\chi^2(2) = 2.32$, $p = .314$), but a significant Test × Group interaction ($\chi^2(2) = 14.26$, $p < .001$). Fixed-effect estimates for the model are presented in Table 4. Post-hoc pairwise comparisons showed first that while the narrative-based group had significantly more words at post-intervention than at pre-intervention ($\beta = 8.77$, $SE = 2.38$, $t(104) = 3.68$, $p < .001$, $d = .83$), the control ($p = .106$) and multimodal narrative-based groups ($p = .834$) did not. It also showed that at post-intervention, only the narrative-based group outperformed the control group ($\beta = -9.44$, $SE = 3.28$, $t(168) = -2.88$, $p = .014$, $d = -.90$); no other group differences were significant at post-intervention (control vs. multimodal narrative-based: $p = .174$; multimodal narrative-based vs. narrative-based: $p = 1.000$) or at pre-intervention (all ps > .30; see Figure 3).

Table 4. Fixed-effect estimates for the linear mixed model predicting TNW

Fixed effect	*β*	*SE*	95% CI	*t*	df	*p*
Intercept	36.612	2.579	[31.56, 41.67]	14.19	123.29	< .001
Condition: Multimodal narrative-based	6.654	3.444	[−0.10, 13.40]	1.93	170.52	.055
Condition: Narrative-based	9.444	3.249	[3.08, 15.81]	2.91	168.13	.004
Test: Pre-intervention	4.100	2.516	[−0.83, 9.03]	1.63	104.00	.106
Condition (Multimodal) × Test (Pre)	−4.645	3.611	[−11.72, 2.43]	−1.29	104.00	.201
Condition (Narrative-based) × Test (Pre)	−12.869	3.466	[−19.66, −6.08]	−3.71	104.00	< .001

Note 1. Reference (baseline) levels: Condition = control group; Test = post-intervention.
Note 2. Model fit: R^2 marginal = .052, R^2 conditional = .502, ICC = .475, AIC = 1730.1, BIC = 1760.4
Note 3. Random effects: Participant intercept variance = 61.67 (*SD* = 7.85); CELF intercept variance = 38.50 (*SD* = 6.21); residual variance = 110.77 (*SD* = 10.53).
Note 4. CIs computed using the Wald method.

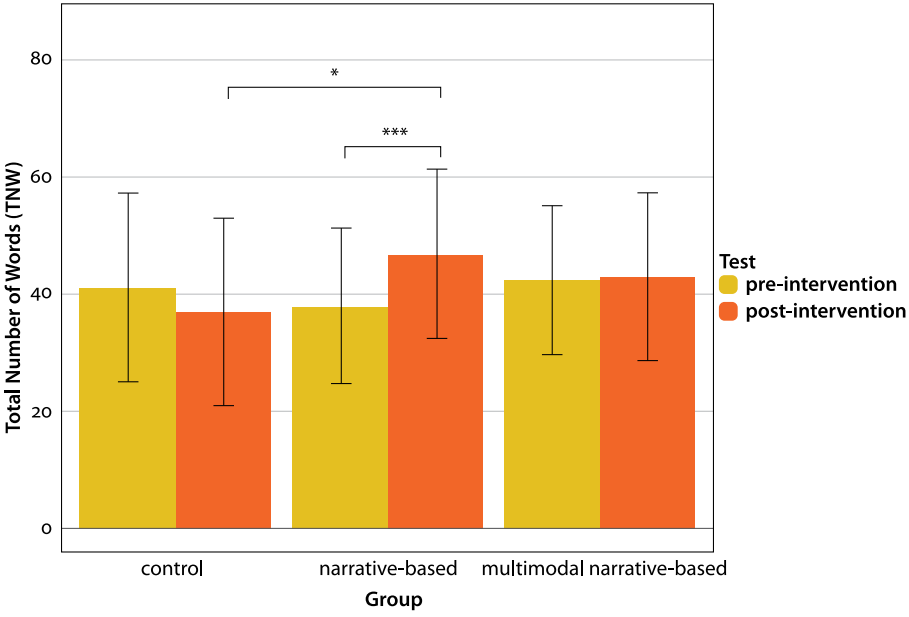


Figure 3. Mean microstructure values (Total Number of Words; TNW) broken down by Test and Group

Note. * $p \leq .05$, *** $p \leq .001$. Error bars represent standard deviations.

Number of Different Words (NDW)

Regarding the NDW measure, the LME model reported no main effects for Test ($\chi^2(1) = 1.52$, $p = .217$) nor Group ($\chi^2(2) = 1.25$, $p = .537$). Only the two-way interaction was found to be significant ($\chi^2(2) = 14.99$, $p < .001$). Fixed-effect estimates for the model are presented in Table 5. Follow-up pairwise comparisons suggested that only the narrative-based group improved from pre- to post-intervention ($\beta = 3.92$, $SE = 1.08$, $t(104) = 3.64$, $p < .001$, $d = .82$), while the control ($p = .074$) and multimodal narrative-based ($p = .897$) groups did not change significantly. It was also observed that the narrative-based group had significantly higher scores than the control group post-intervention ($\beta = -4.41$, $SE = 1.56$, $t(164) = -2.83$, $p = .016$, $d = -.93$); no other group differences reached significance at post-intervention (control vs. multimodal narrative-based: $p = .502$; multimodal narrative-based vs. narrative-based: $p = .546$) or at pre-intervention (all $ps > .88$; see Figure 4).

Table 5. Fixed-effect estimates for the linear mixed model predicting NDW

Fixed effect	*β*	*SE*	95% CI	*t*	df	*p*
Intercept	23.720	1.205	[21.36, 26.08]	19.69	119.82	<.001
Condition: Multimodal narrative-based	2.294	1.633	[−0.91, 5.49]	1.41	165.76	.162
Condition: Narrative-based	4.405	1.543	[1.38, 7.43]	2.85	163.64	.005
Test: Pre-intervention	2.057	1.139	[−0.17, 4.29]	1.81	104.00	.074
Condition (Multimodal) × Test (Pre)	−2.209	1.635	[−5.41, 0.99]	−1.35	104.00	.180
Condition (Narrative-based) × Test (Pre)	−5.980	1.568	[−9.05, −2.91]	−3.81	104.00	<.001

Note 1. Reference (baseline) levels: Condition = control group; Test = post-intervention.
Note 2. Model fit: R^2 marginal = .044, R^2 conditional = .533, ICC = .511, AIC = 1403.6, BIC = 1433.8
Note 3. Random effects: Participant intercept variance = 16.92 (*SD* = 4.11); CELF intercept variance = 6.79 (*SD* = 2.61); residual variance = 22.69 (*SD* = 4.76).
Note 4. CIs computed using the Wald method.

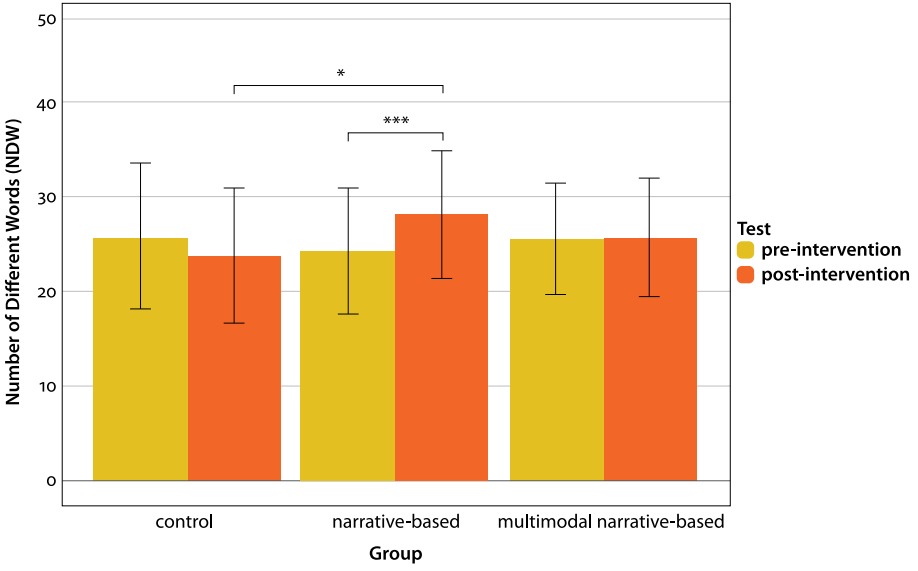


Figure 4. Mean microstructure values (Number of Different Words; NDW) broken down by Test and Group
Note. * $p \leq .05$, *** $p \leq .001$. Error bars represent standard deviations.

Narrative perspective-taking

Regarding the narrative perspective-taking skills, only the main effect of Test was found to be significant ($\chi^2(1) = 4.84$, $p = .028$). Group ($\chi^2(2) = 5.33$, $p = .070$) and the two-way interaction ($\chi^2(2) = 1.96$, $p = .376$) were not significant. Fixed-effect estimates for the model are presented in Table 6. A follow-up comparison indicated that scores were overall higher pre-intervention than post-intervention, regardless of group ($\beta = -0.09$, $SE = 0.04$, $t(112) = -2.27$, $p = .025$, $d = -.30$). No significant baseline group differences were found for narrative perspective-taking (control vs. multimodal narrative-based: $p = .290$; control vs. narrative-based: $p = 1.000$; multimodal narrative-based vs. narrative-based: $p = .082$).

Table 6. Fixed-effect estimates for the linear mixed model predicting narrative perspective-taking

Fixed effect	β	SE	95% CI	t	df	p
Intercept	0.053	0.050	[−0.05, 0.15]	1.04	218.36	.298
Condition: Multimodal narrative-based	0.114	0.072	[−0.03, 0.26]	1.58	218.36	.116
Condition: Narrative-based	0.069	0.070	[−0.07, 0.21]	0.99	218.36	.323
Test: Pre-intervention	0.118	0.065	[−0.01, 0.25]	1.81	112.00	.072
Condition (Multimodal) × Test (Pre)	0.007	0.094	[−0.18, 0.19]	0.07	112.00	.944
Condition (Narrative-based) × Test (Pre)	−0.106	0.091	[−0.28, 0.07]	−1.17	112.00	.244

Note 1. Reference (baseline) levels: Condition = control group; Test = post-intervention.
Note 2. Model fit: R^2 marginal = .049, R^2 conditional = .202, ICC = .161, AIC = 121.9, BIC = 149.4
Note 3. Random effects: Participant intercept variance = 0.016 ($SD = 0.125$); residual variance = 0.081 ($SD = 0.285$).
Note 4. CIs computed using the Wald method.

Exploratory assessment of potential differences between the experimental groups

No significant effects were observed for the multimodal narrative-based group. These findings were not in line with our predictions that the multimodal narrative-based group would show greater improvements than those who receive the same narrative-based intervention without the controlled multimodal component. Therefore, to further understand why our hypothesis was not supported and what was making the multimodal narrative-based group not improve, we conducted a complementary assessment of the implementation of the intervention. This assessment was motivated by our informal observations, while reviewing the video-recorded sessions, that there might be differences between groups in teachers' effort to maintain children's attention and behavior during the sessions.

A research assistant who was blinded to the intervention group watched 18 intervention sessions (i.e., 33% of the sessions) and evaluated children's and teachers' behavior throughout the sessions in terms of (a) the teachers' effort to maintain children's attention and concentration, and (b) the number of times they had to ask children to stay silent, pay attention to the activity, or sit down. On a 1–7 Likert (1 = low effort, no need to ask children to pay attention and 7 = high effort, constantly asking children to pay attention), the results showed the teachers in the multimodal narrative-based group had to intervene way more often ($M = 3.67$, $SD = 1.87$) than those in the narrative-based group ($M = 1.67$, $SD = 0.50$). It was also observed that teachers in the multimodal narrative-based group had to ask for silence, for them to sit down or pay attention an average of 13.67 times per session ($SD = 10.61$), while those in the narrative-based group only an average of 1.44 times per session ($SD = 1.51$).

Narrative skills of children with language and communication needs

We also descriptively assessed whether children with language and communication needs were also able to show improvements by just receiving the tier 1 instruction. From the whole sample, a total of nine children were identified by the schools as having language and communication difficulties (control: $N = 3$; narrative-based: $N = 2$; multimodal narrative-based: $N = 4$).

Due to the small sample and unequal group sizes, no inferential statistics were conducted. Individual data are reported for each child rather than group means on Table 7 (which also includes, for reference, the median scores of children with typical development in the same condition).

Table 7. Individual pre- and post-intervention narrative measure scores for children with language and communication needs, by condition

		MACRO.		MICRO.: TNW		MICRO.: NDW		PERS.	
	Participant (sex, age)	**PRE**	**POST**	**PRE**	**POST**	**PRE**	**POST**	**PRE**	**POST**
Control	Part. A (M, 5;3)	4.5	3.5	44.5	53.5	26	30	0	0
	Part. B (M, 5;6)	2	1.5	25.5	20	21.5	16	0	0
	Part. C (M, 5;6)	0	0.5	–	–	–	–	0	0
	TD	4	4	41.5	33	26.5	23	0	0

Table 7. *(continued)*

		MACRO.		MICRO.: TNW		MICRO.: NDW		PERS.	
	Participant (sex, age)	**PRE**	**POST**	**PRE**	**POST**	**PRE**	**POST**	**PRE**	**POST**
Narrative-based	Part. D (M, 5;11)	0	2	–	–	–	–	0	0
	Part. E (M, 5;5)	4.5	4.5	48	66.5	22	29	0.5	0
	TD	3.5	4.5	35.25	43.75	24.25	28.25	0	0
Multimodal narrative-based	Part. F (M, 5;10)	3.5	4.5	38	54	29.5	32	1	0
	Part. G (M, 5;7)	3.5	5.5	29	38.5	17	23	0.5	0
	Part. H (F, 6;0)	4.5	5	26	47	20.5	27	0	0
	Part. I (F, 5;5)	3	1.5	–	–	–	–	0	0
	TD	4	4	43.5	44	26.75	26.5	0	0

Note 1. MACRO. = macrostructure; MICRO. = microstructure; TNW = total number of words; NDW = number of different words; PERS. = perspective-taking; PRE = pre-intervention; POST = post-intervention; Part. = participant; TD = typical development
Note 2. The TD row shows the median score obtained by children with typical development in the same condition; it is included for reference only and does not represent a directly comparable subgroup.
Note 3. The possible range for narrative macrostructure and narrative perspective-taking was 0–6. The possible range for the narrative microstructure measures was 0–NA, with NA indicating that no theoretical maximum value is defined for this measure.
Note 4. The scores correspond to the average score of the two narrations.
Note 5. A dash indicates that the data was not available.
Note 6. Age is presented in years;months.

Discussion

Effects of the intervention on narrative macrostructure and microstructure

Our results revealed that only the children in the narrative-based significantly improved their narrative macrostructure, and microstructure skills, and that they outperformed the control group not receiving the intervention in the microstructure measures. These findings thus reject our hypothesis that the children receiving the MMN intervention with the specific multimodal component would show greater improvements than those in the narrative-based group.

As just mentioned, the narrative-based group showed significant pre–post improvements in the narrative macrostructure and microstructure measures. We believe that these findings might contribute to the existing evidence on group-based narrative intervention research reporting its potential to boost young children's oral narrative skills (for a review, see Pico et al., 2021). Several factors may account for the significant improvements observed in the narrative-based group relative to the control group. On the one hand, the intervention was designed in a co-creation fashion with a group of more than 90 preschool teachers and speech-language therapists, such that it was tailored to the Catalan context, integrating professionals' usual practices and needs. On the other hand, the MMN intervention program aligns with the narrative intervention principles postulated by Spencer and Petersen (2020), as well as with the strategies commonly used in narrative intervention studies described in the recent reviews (e.g., Favot et al., 2021; Pesco & Gagné, 2017; Pico et al., 2021). These principles and strategies include specific instruction on the macrostructure (i.e., story elements), use of visual support such as pictures or icons, and verbal prompting. The MMN intervention was designed considering previous literature, and therefore includes strategies such as training story elements through a verbal question-and-answer sequence, providing feedback, and using audiovisuals such as video cartoons, icons, and animated pictures. It should be noted that these core intervention features were common to both the narrative-based and multimodal narrative-based conditions, as the two groups only differed in the presence of the systematic multimodal component; as such, these shared design features explain why the MMN intervention as a whole was effective in boosting narrative macrostructure and microstructure relative to no intervention, but do not, on their own, explain the absence of comparable pre–post improvements in the multimodal narrative-based group specifically. We address possible explanations for this differential pattern later in the Discussion.

Effects of the intervention on narrative perspective-taking

The current study not only assessed narrative macrostructure and microstructure, but also narrative perspective-taking, understood as children's ability to talk about the emotions and perspectives of the characters. We observed that the scores for this measure were quite low, ranging from 0–2 when they could go up to 6, and that there were no significant improvements in this measure after the intervention. We believe that even though children are able to recognize and verbalize emotions at the age of three (see Widen & Nelson, 2022, for a review), being able to include them when narrating a story is a more challenging task that is not acquired until later in development, around the age of eight to nine (e.g., Veneziano et al., 2020). Building a narration is complex because it involves

remembering story events, expressing them in order, and using appropriate vocabulary. Therefore, if on top of that, children also need to remember the emotions and perspectives of the characters, which are less salient aspects of the story compared to the macrostructural elements, this makes it even a more complex task. Thus, we believe that preschool children might find it too challenging to talk about the story plot and the characters' emotions and perspectives at the same time. It is also possible that task-related factors contributed to these low scores. For instance, the task may not have been sufficiently explicitly framed to prompt children to reflect on characters' emotions and perspectives (as the target prompt to elicit the story was *Tell me what happened in the story*). Future studies could explore whether using more directly eliciting prompts increases children's spontaneous reference to characters' internal states.

The role of the multimodal component

Complementarily, the present study also sought to assess the effectiveness of the multimodal component within the MMN intervention. Based on the embodied cognition and multimodal enrichment paradigms (Adams, 2010; Foglia & Wilson, 2013; Mathias & von Kriegstein, 2023; see also the reviews by Dargue et al., 2019; Rohlfing, 2019; Vilà-Giménez & Prieto, 2021), we expected significant improvements of the group receiving the MMN intervention with the multimodal component. Nevertheless, our study showed no significant narrative improvements in this group. It is worth noting, however, that the multimodal component tested here combined both children's perception of multimodal cues (via the storyteller and teacher) and children's production of multimodal cues (via embodiment), whereas much of the lab-based evidence motivating our hypotheses was based primarily on gesture perception rather than active production.

There are two possible explanations for these findings, which can be related to the naturalistic nature of the study. First, it should be considered that many studies assessing the causal role of multimodality (particularly gesture) in children's language learning were lab-based studies in which participants were trained individually in a very controlled environment (testing primarily effects of children's perception of gestures). Several factors may make it more difficult to detect the effects of gestures in naturalistic, group-based classroom settings compared to individual lab-based settings. In individual lab-based studies, the experimenter can ensure that each child has full visual access to the gestures produced, that gestures are precisely timed relative to speech, and that the child's attention is directed at the speaker at the moment the gesture is produced. In a classroom setting, by contrast, gestures are directed at the group as a whole, and individual children may vary considerably in their visual access to the teacher, their moment-

to-moment attention, and the extent to which they are distracted by peers or other classroom stimuli. This reduced and more variable visual attention to gestures may weaken their potential benefit for individual children, since the cognitive and memory-related mechanisms through which gestures are thought to support language learning (e.g., strengthening semantic encoding, directing attention to prominent information) depend on the gesture actually being perceived and attended to at the relevant moment (e.g., Aussems & Kita, 2019). As also reported recently by Bernstein et al. (2024), it can therefore be difficult to isolate and test the effects of gestures in natural, group-based settings like classrooms.

Also, a second factor could explain our findings. We thought that the specific way teachers implemented the intervention might have played a key role and for this, a complementary assessment of the implementation of the intervention was conducted. These assessments showed that the classroom dynamics of the multimodal narrative-based group, which required children to repeatedly stand up, enact the stories, and sit down again, may have influenced children's engagement with the narrative content, although the precise mechanism underlying this effect remains unclear. We believe that at least two distinct, non-mutually exclusive explanations can be considered. On the one hand, the physical demands of frequently transitioning between sitting and standing may have imposed a cognitive overload, leaving fewer resources available for processing the narrative's linguistic content (e.g., cognitive load theory; Sweller, 1988). On the other hand, the playful, physically engaging nature of the enactment may instead have distracted children's attention away from the narrative content itself, similar to the "seductive details" effect (Harp & Mayer, 1998), suggesting instead that enactment should be tightly integrated with the narrative as a coherent whole, with bodily actions supporting the development of characters, events, and their temporal and causal relations, rather than representing isolated story fragments.

Another aspect that should be considered is that children had reduced experience with language-based and multimodal interactive activities in the classroom. In some contexts (like ours), children are usually sitting throughout the sessions and are not used to engaging in more active activities such as multimodal story enactment. Overall, our findings suggest that multimodality may be used in classroom settings with preschoolers when it is implemented in a manner that does not interfere with the dynamics of the teacher–group interaction, such as in the narrative-based group, where children are not asked to constantly stand up and sit down to enact the story elements.

In addition to these classroom-based implementation factors, it is also important to consider whether pre-intervention performance and intervention duration may have contributed to the pattern of results observed. Regarding pre-intervention performance, comparisons across studies are difficult given that

different narrative macrostructure coding schemes are used across the literature. However, the studies by Vilà-Giménez and colleagues (2019, 2020) offer a relatively direct point of comparison, as they used a similar age range and an adaptation of the same coding scheme (Demir et al., 2014) as the present study, scored on a comparable 0–6 system. In those studies, pre-test macrostructure scores ranged from approximately 3 to 3.5, closely comparable to the pre-intervention means observed in the present study (ranging from 3.65 to 3.96), suggesting that our groups' baseline performance was not unusually elevated relative to prior literature, and therefore unlikely to fully explain the absence of pre–post improvement in the multimodal narrative-based group through a ceiling or reduced-sensitivity effect alone. Regarding intervention duration, the optimal length needed to detect reliable effects of narrative and multimodal interventions remains an open question in the literature, with effective interventions ranging widely from brief, single-session designs to full-year programs (see Pico et al., 2021, for a review). It therefore remains possible that the nine-session, three-week duration of the present intervention was insufficient to reveal true effects of the multimodal component, particularly given that, as discussed above, children may have required more time and prior experience with multimodal, body-based learning activities before being able to fully benefit from them.

The results of the present study complement and expand on the results of previous investigations. As mentioned above, very few studies have systematically integrated and tested the effects of the multimodality component within a classroom-based narrative intervention, and those that have done so have reported mixed findings. For instance, Pronina et al. (2021) implemented an intervention focusing on embodying mental states through facial expressions. They found that while the intervention was effective in improving children's expressive pragmatic skills, there was no difference between children receiving the embodied and non-embodied intervention. Similarly, Bernstein et al. (2024) compared children's narrative recall after receiving a storybook-reading intervention, with or without drama-based activities involving the teacher and children standing up and acting out as the main characters experiencing the main actions and events of the story plot. The authors found no significant differences between the two groups after the intervention. These two studies incorporated multimodality in a similar fashion as in the present study: involving children embodying the main elements of stories. These findings seem to point out that it is more challenging to isolate the beneficial effects of multimodality when implemented in a more naturalistic context (such as a real classroom setting), potentially due to two distinct, non-mutually exclusive factors: structure and intensity. As mentioned above, frequent, unstructured transitions between sitting and standing may have placed demands on cognitive flexibility, while the physical demands of repeated whole-

body enactment may have imposed an additional working memory load (Sweller, 1988). This aligns with the broader literature on playful versus structured learning contexts, which shows that guided play, in which playful engagement is deliberately structured around specific learning goals, tends to be more effective than more loosely structured play (for a review, see Skene et al., 2022). This suggests that implementing multimodality in a more structured and less intense manner, so that enactment remains tightly integrated with teacher-child communicative interaction, may be more beneficial than a high-intensity or loosely structured approach.

Finally, it is also worth noting that the present study, Bernstein et al. (2024) and Pronina et al. (2021) each assessed different outcome measures (narrative macrostructure/microstructure/perspective-taking, narrative recall/gesture production, and pragmatic/socioemotional skills, respectively), which complicates direct comparison. We therefore believe that future research would benefit from a more consistent battery of outcome measures, as well as more proximal, process-level measures (e.g., children's attention to or spontaneous use of gestures), to clarify whether multimodal input is being processed even when this is not reflected in downstream outcomes.

Limitations and future directions

This intervention study has certain limitations regarding its methodology and classroom implementation. A first limitation of this study is related to the narrative macrostructure coding scheme used: although story complexity was controlled for, the first story featured only one character and the second featured two characters. The coding scheme used in this study did not distinguish whether children introduced one or two characters. Future studies should use coding criteria that can reflect these differences. Additionally, although groups were statistically equivalent at baseline across all narrative outcomes, we observed a non-significant trend toward lower pre-intervention macrostructure scores in the narrative-based group compared to the multimodal narrative-based group; although this difference did not reach significance, it should be considered as a possible contributing factor to the interaction effect observed for this measure. Also, the effects for narrative perspective-taking were small. Given the overall low scores across all groups, care should be taken not to overinterpret the control group's post-intervention score of zero as evidence of a complete absence of this skill. This may reflect a general floor effect or limited sensitivity of the current measurement approach. It would also be interesting to conduct more fine-grained analyses of children's narrative perspective-taking to assess which specific emotional or mental terms children are able to identify and use more frequently.

Regarding the classroom implementation, a full multi-tiered intervention could not be implemented because of the difficulty in coordinating both classroom intervention and small-group or individual interventions. Therefore, these children only received the tier 1 instruction without any extra support. Despite this, as mentioned in the Results section, children with language and communication difficulties still showed some improvements. In addition, we acknowledge that the intervention period was relatively short (only nine sessions over three weeks). A longer intervention might provide more opportunities for children to familiarize themselves with the structure and activities of the intervention, and for teachers to integrate the approach naturally within their curriculum.

Several considerations should be considered related to the multimodal component of the intervention and the complementary assessment of its implementation. First, the complementary assessment of intervention implementation was conducted retrospectively to explore differences across the experimental groups, rather than integrated into the original study design. Relatedly, although the first author reviewed all fidelity evaluations, inter-rater reliability was not calculated for these assessments. Moreover, the *a posteriori* evaluations assessed the extent to which teachers implemented the targeted multimodal strategies, but did not evaluate their naturalness, fluency, or perceived engagement. It is possible that the added demands of enacting the stories made some teachers in the multimodal narrative-based condition appear less natural (e.g., due to overacting or self-consciousness), which could have influenced children's engagement independently of the intended manipulation. Nonetheless, this assessment offered useful insights and directions for improving the integration of multimodality in naturalistic intervention settings in the future. While involving professionals in the design and implementation of the intervention from a participatory approximation to research added ecological validity to the study, the way in which the multimodal component was implemented may have limited its effectiveness. This likely reflects challenges in applying these multimodal strategies in practice. We believe it would have been interesting to first run a pre-study with a single trained interventionist delivering all groups of the intervention to better control implementation and identify any potential problematic issues before the full-scale implementation. We also believe that children may need more prior experience using their bodies to enact the story, especially in contexts where such methods are not typically used. Including drama-based exercises as a preparatory step before introducing a multimodal narrative intervention could help children engage more fully and naturally with the intervention. Finally, to minimize changes in the classroom dynamics and better support children's attention to the language-based interaction, future implementations could either reduce the demands of child enactment (e.g., allocating a dedicated time slot to enact the

full story, or enacting actions and emotions while seated) or, shift the emphasis toward children's perception of multimodal cues rather than requiring active enactment.

Importantly, while the current study is not a definitive test of the isolated effects of multimodality, it does contribute valuable evidence on the general impact of narrative interventions in authentic educational settings and offers guidance for improving the structured implementation of multimodal components in future research. Finally, this study leaves open questions for future research, such as whether children in the multimodal narrative-based condition became more inclined to gesture or embody events post-intervention (as documented in the study by Bernstein et al., 2024), and whether such gestures served a compensatory function for lexical access difficulties.

Implications for practice and conclusions

The findings of this study have implications for classroom intervention practice and research. It was shown that a short nine-session narrative intervention program can be effective in fostering preschool children's oral narrative macrostructure and microstructure skills. Regarding the particular implementation of multimodality in classroom environments, educators should pay attention to the fact that multimodality needs to be carefully and thoughtfully integrated into narrative interventions in a way that it complements and supports communicative interactions.

In sum, we do not interpret the absence of significant effects for the multimodal narrative-based group as evidence that multimodality is ineffective *per se*, but rather as pointing to the need for a more principled implementation of it in classroom contexts. In the present study, and in other classroom-based studies that have incorporated child enactment, embodiment tended to be implemented in a fragmented way, anchored to individual story elements rather than sustained across the narrative. We propose instead that enactment should be integrated with the narrative as a continuous, coherent whole, allowing children to embody the story and its characters throughout, rather than producing isolated gestures tied to single story elements. At the same time, given that our narrative-based condition — which preserved rich, spontaneous multimodal input from teachers and the storyteller without any child embodiment — was the only condition to show significant pre–post improvements, we believe perception-based multimodal input should not be discounted as a substantial driver of narrative learning in classroom settings. We therefore propose that future intervention studies directly compare interventions that combine multimodal perception with continuously integrated embodied production against interventions based solely on

multimodal observation, in order to determine the relative contribution of each to narrative learning in real classrooms.

Funding

This work received funding support by the Spanish Ministry of Science and Innovation, the Spanish State Research Agency and the European Regional Development Fund (ref.: PID2021-123823NB-I00), as well as by the Agency for Management of University and Research Grants (ref.: 2021 SGR_922). The first author would like to acknowledge the financial support of a PhD grant, awarded by the Agency for Management of University and Research Grants (ref.: 2021 FI_B 00778), together with the contract extension awarded by the Department of Translation and Language Sciences at Universitat Pompeu Fabra.
Open Access publication of this article was funded through a Transformative Agreement with Universitat Pompeu Fabra.

Acknowledgements

The authors of this study would like to acknowledge all participating children and their families, as well as schools and teachers. We are also very grateful to Pere Mayans, Imma Canal, and Anna Marco from the Supports and Linguistic Resources Service from the Regional Ministry of Education (Generalitat de Catalunya) for their continuous support in the development of the intervention and in the recruitment of schools. Our gratitude is also extended to Dr. Ĭo Valls-Ratés for acting as storyteller for the intervention materials. Special thanks are given to the research assistants and collaborators that helped collecting the data (Mireia Gómez, Carla Rufí, Claire Luong, Patrick Rohrer, Marco Sanfilippo) and annotate the data (Joel Espejo, Lucía Domingo).

References

Adams, F. (2010). Embodied cognition. *Phenomenology and the cognitive sciences*, *9*, 619–628.

Applebee, A. (1978). *The child's concept of story*. University of Chicago Press.

Aussems, S., & Kita, S. (2019). Seeing iconic gestures while encoding events facilitates children's memory of these events. *Child Development*, *90*(4), 1123–1137.

Babayiğit, S., Roulstone, S. & Wren, Y. (2021). Linguistic comprehension and narrative skills predict reading ability: A 9-year longitudinal study. *British Journal of Educational Psychology*, *91*, 148–168.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48.

Bernstein, K.A., van Huisstede, L., Marley, S.C., Gao, Y.B., Pierce-Rivera, M., Ippolito, E., Restrepo, M.A., Millinger, J., Brantley, K., & Gantwerker, J. (2024). Gesture like a kitten and you won't forget your tale: Drama-based, embodied story time supports preschoolers' narrative skills. *Early Childhood Research Quarterly*, *66*, 178–190.

Bolk, E., Scheper, A., Lissenberg-Witte, B., Goverts, T., & Kramer, S.E. (2026). Development of narrative abilities in (re)telling a story in typically developing Dutch school-aged children. *Journal of Speech, Language, and Hearing Research*.

Botting, N. (2002). Narrative as a tool for the assessment of linguistic and pragmatic impairments. *Child Language Teaching and Therapy*, *18*(1), 1–21.

Corbett, B.A., Key, A.P., Qualls, L., Fecteau, S., Newsom, C., Coke, C., & Yoder, P. (2016). Improvement in social competence using a randomized trial of a theatre intervention for children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *46*(2), 658–672.

Daelman, J., Alighieri, C., Van Lierde, K., Simon, E., Altinkamis, F., Baudonck, N., & D'haeseleer, E. (2024). Effect of a group-based narrative language training in typically developing Turkish-Dutch children. *Folia Phoniatrica et Logopaedica*, *76*(2), 192–205.

Dargue, N., Sweller, N., & Jones, M.P. (2019). When our hands help us understand: A meta-analysis into the effects of gesture on comprehension. *Psychological Bulletin*, *145*(8), 765–784.

Demir, Ö. E., Fisher, J.A., Goldin-Meadow, S., & Levine, S.C. (2014). Narrative processing in typically developing children and children with early unilateral brain injury: Seeing gesture matters. *Developmental Psychology*, *50*(3), 815–828.

Demir, Ö. E., Levine, S.C., & Goldin-Meadow, S. (2015). A tale of two hands: Children's early gesture use in narrative production predicts later narrative structure in speech. *Journal of Child Language*, *42*(3), 662–681.

Dickinson, D.K., & McCabe, A. (2001). Bringing it all together: The multiple origins, skills, and environmental supports of early literacy. *Learning Disabilities Research & Practice*, *16*(4), 186–202.

Dodd, J.L., Ocampo, A., & Kennedy, K.S. (2011). Perspective taking through narratives: An intervention for students with ASD. *Communication Disorders Quarterly*, *33*(1), 23–33.

Donolato, E., Toffalini, E., Rogde, K., Nordahl-Hansen, A., Lervåg, A., Norbury, C., & Melby-Lervåg, M. (2023). Oral language interventions can improve language outcomes in children with neurodevelopmental disorders: A systematic review and meta-analysis. *Campbell Systematic Reviews*, *19*(4), e1368.

Favot, K., Carter, M., & Stephenson, J. (2021). The effects of oral narrative intervention on the narratives of children with language disorder: A systematic literature review. *Journal of Developmental and Physical Disabilities*, *33*(4), 489–536.

Florit-Pons, J., Igualada, A., & Prieto, P. (2025). Evaluating the feasibility and preliminary effectiveness of a multi-tiered multimodal narrative intervention program for preschool children. *Language, Speech, and Hearing Services in Schools*, *56*(1), 17–41.

Florit-Pons, J., Prieto, P., & Igualada, A. (2025). Co-creation of a narrative intervention program for speech–language pathology and educational settings. *American Journal of Speech-Language Pathology*, *34*(6), 3279–3299.

Foglia, L., & Wilson, R.A. (2013). Embodied cognition. *WIREs Cognitive Science*, *4*, 319–325.

Fox, J., & Weisberg, S. (2019). *An R companion to applied regression* (3rd ed.). Sage.

Frydman, J. S., & Mayor, C. (2024). A scoping review on the use and potential of school-based drama therapy to enhance socio-emotional skills in early childhood. *Early Childhood Education Journal*, *52*(4), 669–680.

Gillam, S. L., Gillam, R. B., & Laing, C. E. (2014). *Supporting knowledge in language and literacy*. Utah State University. https://meadowscenter.org/project/supporting-knowledge-in-language-and-literacy-skill/

Gillam, S. L., Olszewski, A., Fargo, J., & Gillam, R. B. (2014). Classroom-based narrative and vocabulary instruction: Results of an early-stage, nonrandomized comparison study. *Language, Speech, and Hearing Services in Schools*, *45*(3), 204–219.

Glenberg, A. M. (2010). Embodiment as a unifying perspective for psychology. *WIREs Cognitive Science*, *1*(4), 586–596.

Guilbert, D., Sweller, N., & Van Bergen, P. (2021). Emotion and gesture effects on narrative recall in young children and adults. *Applied Cognitive Psychology*, *35*(4), 873–889.

Harp, S. F., & Mayer, R. E. (1998). How seductive details do their damage: A theory of cognitive interest in science learning. *Journal of Educational Psychology*, *90*(3), 414–434.

Henry, A. R., & Solari, E. J. (2020). Targeting oral language and listening comprehension development for students with autism spectrum disorder: A school-based pilot study. *Journal of Autism and Developmental Disorders*, *50*(10), 3763–3776.

Hettiarachchi, S. (2022). The effectiveness of the adapted Box Clever language intervention programme in the development of vocabulary and narrative skills of deaf and hard of hearing children. *Deafness & Education International*, *24*(1), 65–91.

Hostetter, A. B., & Alibali, M. W. (2008). Visible embodiment: Gestures as simulated action. *Psychonomic Bulletin & Review*, *15*(3), 495–514.

Hostetter, A. B., Bieda, K., Alibali, M. W., Nathan, M. J., & Knuth, E. J. (2006). Don't just tell them, show them! Teachers can intentionally alter their instructional gestures. In R. Sun (Ed.), *Proceedings of the 28th Annual conference of the cognitive science society* (pp. 1523–1528). Erlbaum.

Ionescu, T., & Ilie, A. (2018). Language learning in preschool children: An embodied learning account. *Early Child Development and Care*, *188*(1), 4–15.

Justice, L. M., Bowles, R. P., Kaderavek, J. N., Ukrainetz, T. A., Eisenberg, S. L., & Gillam, R. B. (2006). The index of narrative microstructure: A clinical tool for analyzing school-age children's narrative performances. *American Journal of Speech-Language Pathology*, *15*, 177–191.

Kaufman, A. S., & Kaufman, N. L. (1990). *Kaufman Brief Intelligence Test*. Pearson.

Kirk, E., Donnelly, S., Furman, R., Warmington, M., Glanville, J., & Eggleston, A. (2022). The relationship between infant pointing and language development: A meta-analytic review. *Developmental Review*, *64*, 101023.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1–26.

Lenth, R. V. (2021). *emmeans: Estimated marginal means, aka least-squares means* (R Package Version 1.5.5–1). https://cran.r-project.org/web/packages/emmeans/emmeans.pdf

Lüdecke, D., Ben-Shachar, M., Patil, I., Waggoner, P., Makowski, D. (2021). performance: An R package for assessment, comparison and testing of statistical models. *Journal of Open Source Software*, *6*(60), 3139.

Macoun, A., & Sweller, N. (2016). Listening and watching: The effects of observing gesture on preschoolers' narrative comprehension. *Cognitive Development*, *40*, 68–81.

MacWhinney, B. (2000). *The CHILDES Project: Tools for Analyzing Talk* (3rd ed.). Lawrence Erlbaum Associates.

Mathew, M., Yuen, I., & Demuth, K. (2018). Talking to the beat: Six-year- olds' use of stroke-defined non-referential gestures. *First Language*, *38*, 111–128.

Mathias, B., & von Kriegstein, K. (2023). Enriched learning: Behavior, brain, and computation. *Trends in Cognitive Sciences*, *27*(1), 81–97.

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. University of Chicago Press.

Nelson, K.E., & Khan, K.S. (2019). A dynamic systems account incorporating eight narrative developmental stages: Chapter 8. New frontiers in facilitating narrative skills in children and adolescents. In E. Veneziano & A. Nicolopoulou (Eds.), *Narrative, Literacy and Other Skills: Studies in intervention* (pp. 173–199). John Benjamins Publishing Company.

Nicolopoulou, A., Schnabel, K., Ilgaz, H., Brockmeyer, C., & De Sá, A.B. (2015). Using a narrative- and play-based activity to promote low-income preschoolers' oral language, emergent literacy, and social competence. *Early Childhood Research Quarterly*, *31*, 147–162.

Noble, C., Sala, G., Peter, M., Lingwood, J., Rowland, C., Gobet, F., & Pine, J. (2019). The impact of shared book reading on children's language skills: A meta-analysis. *Educational Research Review*, *28*, 100290.

Perniss, P. (2018). Why we should study multimodal language. *Frontiers in Psychology*, *9*, 1109.

Pesco, D., & Gagné, A. (2017). Scaffolding narrative skills: A meta-analysis of instruction in early childhood settings. *Early Education and Development*, *28*(7), 773–793.

Petersen, D.B., Staskowski, M., Spencer, T.D., Foster, M.E., & Brough, M.P. (2022). The effects of a multi-tiered system of language support on kindergarten oral and written language: A large-scale randomized controlled trial. *Language, Speech, and Hearing Services in Schools*, *53*(1), 44–68.

Pico, D.L., Hessling Prahl, A., Biel, C.H., Peterson, A.K., Biel, E.J., Woods, C., & Contesse, V.A. (2021). Interventions designed to improve narrative language in school-age children: A systematic review with meta-analyses. *Language, Speech, and Hearing Services in Schools*, *52*(4), 1109–1126.

Prieto, P., Cravotta, A., Kushch, O., Rohrer, P., & Vilà-Giménez, I. (2018). Deconstructing beat gestures: A labelling proposal. In K. Klessa, J. Bachan, A. Wagner, M. Karpiński, & D. Śledziński (Eds.), *Proceedings of the 9th International Conference on Speech Prosody* (pp. 201–205). Poznań, Poland.

Pronina, M., Hübscher, I., Holler, J., & Prieto, P. (2021). Interactional training interventions boost children's expressive pragmatic abilities: Evidence from a novel multidimensional testing approach. *Cognitive Development*, *57*, 101003.

Roberts, B.R.T., MacLeod, C.M., & Fernandes, M.A. (2022). The enactment effect: A systematic review and meta-analysis of behavioral, neuroimaging, and patient studies. *Psychological Bulletin*, *148*(5–6), 397–434.

Rohlfing, K. (2019). Learning language from the use of gestures. In J.S. Horst & J.V. Torkildsen (Eds.), *International Handbook of Language Acquisition* (pp. 213–233). Routledge.

Rohrer, P. L., Florit-Pons, J., Vilà-Giménez, I., & Prieto, P. (2022). Children use non-referential gestures in narrative speech to mark discourse elements which update common ground. *Frontiers in Psychology*, *12*, 661339.

Schapira, R., & Grazzani, I. (2025). Shared book reading and promoting social and emotional competences in educational settings: A narrative review. *Frontiers in Psychology*, *16*, 1622536.

Seccia, A., & Goldin-Meadow, S. (2024). Gestures can help children learn mathematics: How researchers can work with teachers to make gesture studies applicable to classrooms. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *379*(1911), 20230156.

Skene, K., O'Farrelly, C. M., Byrne, E. M., Kirby, N., Stevens, E. C., & Ramchandani, P. G. (2022). Can guidance during play enhance children's learning and development in educational contexts? A systematic review and meta-analysis. *Child Development*, *93*(4), 1162–1180.

So, W. C., Sim Chen-Hui, C., & Low Wei-Shan, J. (2012). Mnemonic effect of iconic gesture and beat gesture in adults and children: Is meaning in gesture important for memory recall? *Language and Cognitive Processes*, *27*(5), 665–681.

Spencer, T. D., & Petersen, D. B. (2020). Narrative intervention: Principles to practice. *Language, Speech, and Hearing Services in Schools*, *51*(4), 1081–1096.

Spencer, T. D., & Petersen, D. B. (2012). *Story Champs® 2.0*. Language Dynamics Group. https://www.languagedynamicsgroup.com/story-champs-2/story-champs-about/

Spencer, T. D., & Petersen, D. B. (2018). Narratives bridge oral and written language: An oral language intervention study with writing outcomes. *Language, Speech, and Hearing Services in Schools*, *49*, 569–581.

Spencer, T. D., Weddle, S. A., Petersen, D. B., & Adams, J. A. (2018). Multi-tiered narrative intervention for preschoolers: A Head Start implementation study. *NHSA Dialog*, *20*(1), 1–28.

Suggate, S. P., Lenhart, J., Vaahtoranta, E., & Lenhard, W. (2021). Interactive elaborative storytelling fosters vocabulary in preschoolers compared to repeated-reading and phonemic awareness interventions. *Cognitive Development*, *57*, 100996.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, *12*(2), 257–285.

Veneziano, E., Le Normand, M.-T., Plumet, M.-H., & Elie-Deschamps, J. (2020). Promoting narrative skills in 5- to 8-year-old French-speaking children: The effects of a short conversational intervention. *First Language*, *40*(3), 225–250.

Vilà-Giménez, I., Igualada, A., & Prieto, P. (2019). Observing storytellers who use rhythmic beat gestures improves children's narrative discourse performance. *Developmental Psychology*, *55*(2), 250–262.

Vilà-Giménez, I., & Prieto, P. (2021). The value of non-referential gestures: A systematic review of their cognitive and linguistic effects in children's language development. *Children*, *8*(2), 148.

Vilà-Giménez, I., & Prieto, P. (2020). Encouraging kids to beat: Children's beat gesture production boosts their narrative performance. *Developmental Science*, *23*(6), 1–14.

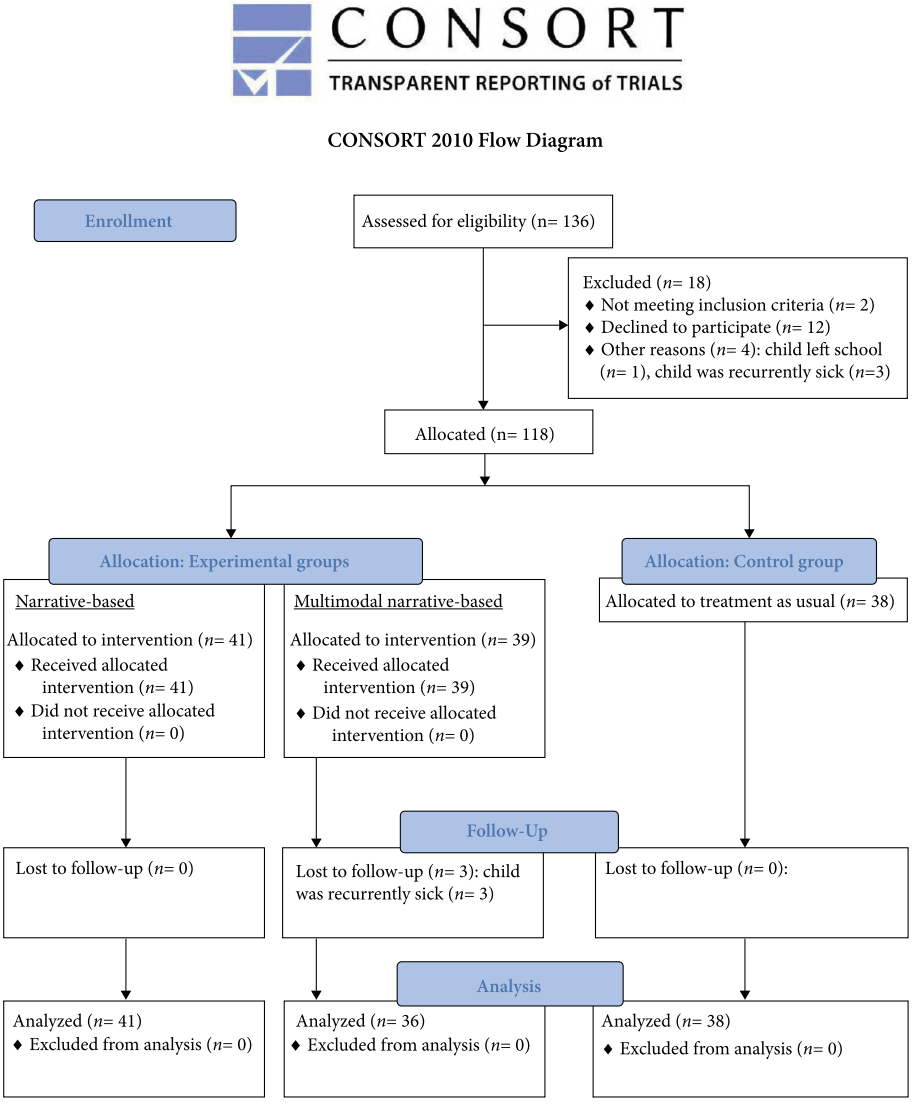
West, G., Snowling, M. J., Lervåg, A., Buchanan-Worster, E., Duta, M., Hall, A., McLachlan, H., & Hulme, C. (2021). Early language screening and intervention can be delivered successfully at scale: Evidence from a cluster randomized controlled trial. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *62*(12), 1425–1434.

Widen, S. C., & Nelson, N. L. (2022). Differentiation and language acquisition in children's understanding of emotion. In D. Dukes, A. C. Samson & E. A. Walle (Eds.), *The Oxford handbook of emotional development* (pp. 174–187). Oxford Academic.

Wiig, E. H., Semel, E., & Secord, W. A. (2013). *Clinical Evaluation of Language Fundamentals* | Fifth edition *(CELF-5)*. Pearson. https://www.pearsonassessments.com/en-us/Store/Professional-Assessments/Speech-%26-Language/Clinical-Evaluation-of-Language-Fundamentals-%7C-Fifth-Edition/p/100000705?srsltid=AYCX9vKVot-UwDV3LvTYvbZ_Msta079iLgDqKadvDwPFRwjInfE5Pcxv

Appendices

Appendix I. *Consort flow diagram*



Appendix II. *Narrative macrostructure coding scheme (adapted from Demir et al., 2014)*

Score	Description
0	The story does not include any descriptive sequence.
1	The story includes one descriptive sequence (without any temporal sequence).
2	The story includes an action sequence (such as the main character and the problem).
3	The child produces an incomplete narrative that lacks two or more of the macrostructure elements (character, problem, attempt, solution, final).
4	The child produces an incomplete narrative that lacks one of the macrostructure elements (character, problem, attempt, solution, final).
5	The child produces a complete narrative that includes all macrostructure elements.
6	The child produces a complete narrative that includes all macrostructure elements and also adds details about the story.

Appendix III. *Narrative perspective-taking coding scheme (adapted from Dodd et al., 2011)*

Score	Description
0	The story does not include any emotion.
1	The story includes one emotion (such as *angry, happy, sad, surprised*).
2	The story includes two or more emotions (such as *angry, happy, sad, surprised*).
3	The story includes one emotion (such as *angry, happy, sad, surprised*) and its cause.
4	The story includes two or more emotions (such as *angry, happy, sad, surprised*) and the cause of at least two emotions.
+1	The story includes one mental term (such as *thinking, realizing, willing, wanting*).
+2	The story includes two or more mental terms (such as *thinking, realizing, willing, wanting*).

Appendix IV. *Example of narrative macrostructure and perspective-taking coding: Target elements and a sample child story*

Story and main story elements	Sample child story
Story with one character Story elements: – Character (C): The mouse. – Problem (P): The wind blows the socks off the clothesline because the mouse does not have any clothespins. – Attempt (A): The mouse hangs up the socks again (but the wind constantly blows them off the clothesline). – Solution (S): As the socks have holes, the mouse decides to hang them up by passing them through the clothesline. – End (E): The wind blows again and, this time, the socks do not fall down. Perspective-taking elements: – The mouse is happy while hanging up the socks. – The mouse is angry because the wind constantly blows the socks off. – The mouse thinks about how to solve the problem. – The mouse has the idea to pass the socks through the clothesline. – The mouse is satisfied because the wind blows and does not blow the socks off.	**Original version (Catalan)** Que un ratolí (C) estava posant la roba en l'estenedor. I quan la posava, després l'aire li treia la troba (P). La tornava a penjar i li tornava a volar (A). Al final ha agafat el fil i hi ha posat la roba per dins (S). I així el vent després no li feia volar més la roba (E). **Translated version (English)** A mouse (C) was hanging clothes on the clothesline. And every time he hung them up, the wind would blow them off (P). He hung them again, and they would fly off again (A). Finally, he takes the clothesline and passes the clothes through it (S). That way, the wind could not blow the clothes off anymore (E).

Note. Example based on the *Löchersocken* story (*Die Sendung mit der Maus*), available at https://www.wdrmaus.de/filme/mausspots/loechersocken.php5

Appendix V. *Storyteller scripts for the three stories used during the MMN intervention*

Story	Script
STORY 1	**Original version (Catalan)** *En Meloix i la banyera* El capibara Meloix està jugant amb la seva pilota i s'ho passa súper bé, està molt feliç. De sobte, la pilota cau en una bassa de fang i embruta en Meloix. Quina cara més bruta que li ha quedat! En Meloix s'enfada perquè està ben enfangat i pensa què pot fer per tornar a estar net. En Meloix té una idea i busca dins l'armari alguna cosa per netejar-se. Treu totes les joguines de l'armari fins que troba una tovallola, un raspall i sabó. Agafa les coses i se'n va cap a la banyera. En Meloix prepara la banyera i s'hi posa a dins fent un salt. Primer de tot neteja la pilota i després es renta tot el cos. Al final, en Meloix està molt content perquè ja està ben net. I torna a jugar amb la pilota! **Translated version (English)** The capybara Meloix is playing with his ball and having a great time; he's really happy. Suddenly, the ball falls into a mud puddle and gets Meloix all dirty. What a dirty face he ends up with! Meloix gets upset because he's covered in mud and thinks about what he can do to get clean again. Meloix has an idea and looks in the closet for something to clean himself with. He takes all the toys out of the closet until he finds a towel, a brush, and soap. He grabs everything and heads to the bathtub. Meloix fills the bathtub and jumps right in. First he cleans the ball, and then he cleans his whole body. In the end, Meloix is very happy because now he's clean again. And he goes back to playing with his ball!
STORY 2	**Original version (Catalan)** *En Meloix i els plàtans* El capibara Meloix està caminant tranquil·lament. De sobte li comença a fer soroll la panxa, té molta i molta gana! Per sort, al seu davant troba un arbre amb molts plàtans grocs i molt llaminers, i vol agafar-ne un! Els plàtans, però, són molt amunt i en Meloix s'adona que no hi arriba. En Meloix pensa què pot fer per agafar els plàtans i té una idea! Porta una caixa de fusta per acostar-se als plàtans, la posa a sota dels plàtans i hi puja a sobre. En Meloix comença a saltar per agafar els plàtans, però així tampoc no hi arriba. En Meloix es posa trist. Pensa, pensa i té una altra idea! En Meloix agafa una pedra gran i la posa a sobre de la caixa de fusta. Però la posa a l'inrevés! Quan hi puja, en Meloix trontolla i s'espanta molt, però no cau i altra vegada ai ai ai trontolla i *patapum*, en Meloix cau a terra! Com que veu que així no hi arriba, decideix enfilar-se a l'arbre. Quan és a dalt i està a punt d'agafar els plàtans, *patapum*, torna a caure a terra. Es posa trist, és clar, s'ha fet mal i no arriba als plàtans. En aquell mateix

Appendix V. *(continued)*

Story	Script
	moment apareix el mico, en Llampoc, un molt bon amic d'en Meloix, i veu que en Meloix ha caigut i està molt trist i decebut. El mico Llampoc veu que vol agafar els plàtans però no pot agafar-los, és per això que en Meloix està trist. Com que és el seu amic i el vol ajudar, ràpidament s'enfila a sobre d'en Meloix i agafa els plàtans grocs. Al final, els dos mengen cadascun un plàtan ben contents! **Translated version (English)** *Meloix and the bananas* The capybara Meloix is walking along calmly. Suddenly his tummy starts to rumble, he's really, really hungry! Luckily, right in front of him he finds a tree full of lots of yellow, delicious-looking bananas, and he wants to grab one! The bananas, though, are very high up, and Meloix realizes he can't reach them. Meloix thinks about what he can do to get the bananas, and he has an idea! He brings over a wooden box to get closer to the bananas, puts it under the tree, and climbs on top of it. Meloix starts jumping to grab the bananas, but he still can't reach them that way either. Meloix feels sad. He thinks and thinks, and has another idea! Meloix picks up a big rock and puts it on top of the wooden box. But he puts it on upside down! When he climbs up, Meloix wobbles and gets really scared, but he doesn't fall; and then, oh no, he wobbles again, and *boom*, Meloix falls to the ground! Since he sees that he can't reach them that way, he decides to climb up the tree instead. When he's up at the top and just about to grab the bananas, *boom*, he falls to the ground again. He feels sad, of course; he's hurt himself and still can't reach the bananas. At that very moment, along comes the monkey, Llampoc, a very good friend of Meloix's, and he sees that Meloix has fallen and is very sad and disappointed. The monkey Llampoc sees that Meloix wants to grab the bananas but can't reach them, and that's why he's sad. Since he's his friend and wants to help him, he quickly climbs up onto Meloix and grabs the yellow bananas. In the end, the two of them each happily eat a banana!
STORY 3	**Original version (Catalan)** *En Meloix i l'autobús* El capibara Meloix vol anar a la platja en autobús. Té moltes ganes d'anar a la platja, per això se'n va a la parada de l'autobús ben content, i s'asseu a esperar que arribi. Al cap d'una estona, en Meloix veu que l'autobús s'acosta, així que s'aixeca i mou el braç per cridar l'autobús i que s'aturi. *Ups*, no l'ha vist! En Llampoc, el seu amic mico que condueix l'autobús, ha passat de llarg i no s'ha aturat a recollir en Meloix! En Meloix està ben sorprès que en Llampoc no

Appendix V. *(continued)*

Story	Script
	s'hagi aturat. Quan en Llampoc s'adona que no s'ha aturat, para de cop l'autobús i fa senyals a en Meloix perquè hi pugi. En Meloix va corrents cap a l'autobús, posa el seu flotador de platja a sobre de l'autobús i hi puja ben cansat. A la següent parada, la gallineta Zepa també està esperant l'autobús, però una altra vegada, el mico Llampoc passa de llarg! La gallineta Zepa comença a córrer i a empaitar l'autobús. En Llampoc ho veu i para de cop. La gallineta Zepa està molt enfadada amb en Llampoc perquè no s'ha aturat a recollir-la. En Meloix, però, l'ajuda a pujar a l'autobús. Quan són a mig camí, de sobte, l'autobús comença a anar lent i s'atura de cop! En Meloix, la Zepa i en Llampoc baixen de l'autobús i van a veure què ha passat. Se'ls ha punxat una roda! El mico Llampoc i la gallineta Zepa estan molt preocupats, no saben com solucionar el problema. En Meloix pensa què pot fer per arreglar la roda, i té una idea! Agafa el seu flotador de platja que té forma de roda i el posa al lloc de la roda foradada. Al final, l'autobús pot tornar a marxar i els tres se'n van ben contents cap a la platja!
	Translated version (English) The capybara Meloix wants to go to the beach by bus. He's really looking forward to it, so he happily heads to the bus stop and sits down to wait for it to arrive. After a while, Meloix sees the bus coming, so he stands up and waves his arm to call the bus over and get it to stop. *Oops*, he didn't see him! Llampoc, his monkey friend who drives the bus, has driven right past without stopping to pick Meloix up! Meloix is really surprised that Llampoc didn't stop. When Llampoc realizes he didn't stop, he slams on the brakes and waves for Meloix to hop on. Meloix runs over to the bus, puts his beach float on top of the bus, and climbs aboard, exhausted. At the next stop, the little hen Zepa is also waiting for the bus, but once again, the monkey Llampoc drives right past! Zepa starts running and chasing after the bus. Llampoc sees her and stops short. Zepa is very angry with Llampoc for not stopping to pick her up. Meloix, though, helps her climb onto the bus. When they're halfway there, suddenly the bus starts going slower and slower and comes to a sudden stop! Meloix, Zepa, and Llampoc get off the bus to see what's happened. They've got a flat tyre! The monkey Llampoc and the hen Zepa are very worried: they don't know how to fix the problem. Meloix thinks about what he can do to fix the tyre, and he has an idea! He takes his beach float, which is shaped like a wheel, and puts it in place of the flat tyre. In the end, the bus can set off again, and the three of them head happily off to the beach!

Address for correspondence

Júlia Florit-Pons
Department of Applied Pedagogy and Educational Psychology
Universitat de les Illes Balears
Spain
julia.florit@upf.edu
https://orcid.org/0000-0002-9025-5239

Co-author information

Alfonso Igualada
NeuroDevelop eHealth Lab, eHealth Center and Department of Psychology and Education Sciences
Universitat Oberta de Catalunya
aigualada@uoc.edu

Pilar Prieto
Institució Catalana de Recerca i Estudis Avançats
Universitat Pompeu Fabra
pilar.prieto@upf.edu

Publication history

Date received: 17 March 2026
Date accepted: 21 October 2026
Published online: 19 November 2026